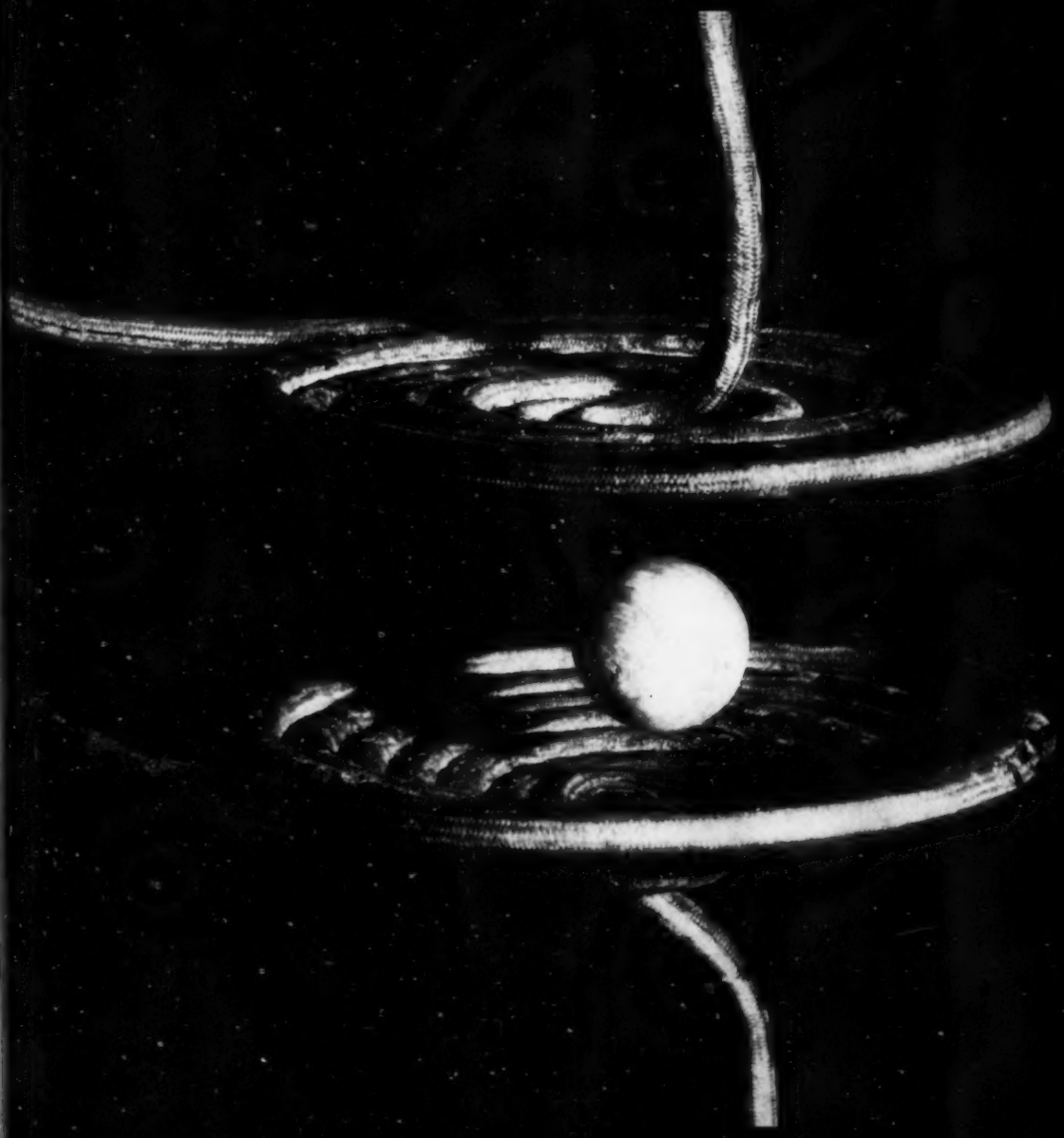
JOURNAL OF APPLIED PHYSICS



JOURNAL OF APPLIED PHYSICS

Vol. 23, No. 5 MAY, 1952

In This Issue

		Page
Dielectric Constant of Dry Air	J. V. Hughes and H. L. Armstrong	501
Computation of Dielectric Constants	R. M. Redheffer, R. C. Wildman, and V. O'Gorman	505
Contribution to the Theory of the Silicon Carbide Contact	Else Holm	509
Initiation of Electrical Breakdown in Vacuum	Lawrence Cranberg	518
Eddy-Currents in Solid Cylindrical Cores Having Non-Uniform Permeability	Harold Aspden	523
Information Theory and Its Application to Taxonomy	D. K. C. MacDonald	529
Theory of Jet Formation by Charges with Lined Conical Cavities	Emerson M. Pugh, R. J. Eichelberger, and Norman Rostoker	532
Experimental Verification of the Theory of Jet Formation by Charges with Lined Conical Cavities	R. J. Eichelberger and Emerson M. Pugh	537
Condition for the Creeping of a Liquid Past a Mercury Seal	R. S. Jessup	543
Electromagnetic Levitation of Solid and Molten Metals	E. C. Okress, D. M. Wroughton, G. Comenetz, P. H. Brace, and J. C. R. Kelly	545
Arcing at Electrical Contacts on Closure. Part III. Development of an Arc	L. H. Germer and J. L. Smith	553
Instabilities in the Smooth-Anode Cylindrical Magnetron	Lawrence A. Harris	562
Effect of Velocity Distribution in a Modulated Electron Stream	Dean A. Watkins	568
Magnetically Induced Ultrasonic Velocity Changes in Polycrystalline Nickel	S. J. Johnson and T. F. Rogers	574
Heat Conduction Alloys at Low Temperatures	I. Estermann and J. E. Zimmerman	578
Radiation Pattern of an Antenna over a Circular Ground Screen	James E. Storer	588
Drift Velocity of Electrons in Argon and Argon Mixtures	J. M. Kirshner and D. S. Toffolo	594

Announcements:

ASME Symposium on Shock and Vibration Instrumentation Planned for 1952	504
Conference on Medium Energy Nuclear Physics	522
Summer Courses	536

Letters to the Editor:

Thermionic Emission and Electrical Conductivity of Oxide-Coated Cathodes	Shin-Ichiro Narita	599
Errata: General Theory of Electromagnetic Horns	A. F. Stevenson	599
Erratum: The Thermal Conductivity of Metals at High Temperature	C. L. Hogan and R. B. Sawyer	600
Erratum: Review of Book Entitled "Microwave Electronics"	F. Carl Willenbrock	600
Application of Nutting's Equation to the Viscoelastic Behavior of Certain Polymeric Systems	G. W. Scott Blair	600
Application of Nutting's Equation to the Viscoelastic Behavior of Certain Polymeric Systems	Rolf Buchdahl and Lawrence E. Nielsen	600
X-Ray Line Broadening of Plastically Deformed Marble	D. Rosenthal and M. Kaufman	600
Books Reviewed		602
Other Books Received		604
Index to Advertisers		xxiv

BOARD OF EDITORS

Elmer Hutchisson, *Editor*

Earle C. Gregg, *Assistant Editor*

Associate Editors

Howard H. Aiken
Ray E. Bolz
Lyle D. Borst
Robert D. Heidenreich
James Hillier
Leonard B. Loeb
W. James Lyons
Eric Reissner
Roman Smoluchowski

Publication of the American Institute of Physics

OFFICERS

George R. Harrison, *Chairman*
George B. Pegram, *Treasurer*

GOVERNING BOARD

Karl K. Darrow
G. J. Dienes
S. A. Goudsmit
G. P. Harnwell
George R. Harrison
F. V. Hunt
Rudolph Kingslake
Hugh S. Knowles
C. C. Lauritsen
F. W. Loomis
William F. Meggers
George B. Pegram
I. I. Rabi
Duane Roller
Frederick Seitz
Francis G. Slack
John C. Steinberg
Richard M. Sutton
Mark W. Zemansky

Consultant on Printing

Melvin Loos

ADMINISTRATIVE STAFF

Henry A. Barton, *Director*
Wallace Waterfall, *Executive Secretary*
Theodore Vorburger, *Advertising Manager*
Ruth F. Bryans, *Publication Manager*
Betty H. Goodfriend, *Assistant Secretary*
Eileen B. Neuberger, *Circulation Manager*
Kathryn Setze *Assistant Treasurer*

The Journal of Applied Physics, published monthly at Prince and Lemon Sts., Lancaster, Pa., is devoted to physics in its role as the science basic to other natural sciences and to the arts and industries. Previous to 1937 this publication was known as Physics. The Journal of Applied Physics publishes editorials and reviews, as well as technical papers of applied physics. Articles appearing in it are indexed regularly, according to their field, by Physics Abstracts, Chemical Abstracts, Engineering Index, Inc., Industrial Arts Index, etc.

Manuscripts should be submitted to Elmer Hutchisson, Editor, Case Institute of Technology, Cleveland 6, Ohio. Unless otherwise stated, submission of a manuscript is a representation that it has been neither copyrighted, published, nor submitted for publication elsewhere. The authors' institutions are requested to pay a **publication charge** of $8 per page which, if honored, entitles them to 100 reprints without further charge. Instructions will be sent with galley proofs.

Subscriptions, renewals, and orders for back numbers should be addressed to the American Institute of Physics, 57 East 55 Street, New York 22, New York.

Subscription Price	U.S. and Canada	Elsewhere
To members of the American Institute of Physics and Affiliated Societies	$10.00	$11.00
To all others	12.00	13.00

Back Number Prices

Yearly back number rate when complete year is available: $10.00. Single copies: $1.00 each.

Changes of address should be sent to the American Institute of Physics; **corrected proofs** should be sent to the Publication Manager, American Institute of Physics. **Advertising rates** supplied on request. Orders, advertising copy, and cuts should be sent to the American Institute of Physics.

Entered as second class matter January 22, 1937, at the Post Office at Lancaster, Pennsylvania, under the Act of March 3, 1879. Accepted for mailing at the special rate of postage provided for in paragraph (d-2), section 34.40, P. L. & R. of 1948, authorized May 2, 1932.

Journal
of
Applied Physics

Volume 23, Number 5 May, 1952

The Dielectric Constant of Dry Air

J. V. Hughes* and H. L. Armstrong†
Queen's University, Kingston, Ontario, Canada
(Received December 3, 1951)

A sensitive and accurate method for determining the real part of the dielectric constant of gases at microwave frequencies is described. The method eliminates some possible sources of systematic error inherent in earlier methods.

The dielectric constant of dry air at N.T.P. has been determined as 1.000569, at a frequency of 3000 Mc/sec. This is in good agreement with other results, at 24,000 Mc/sec, at 9000 Mc/sec, and at low frequencies, thus verifying that the dielectric constant is independent of frequency up to nearly 30,000 Mc/sec An earlier result suggesting the contrary was found to be in error because of insufficient drying of the air.

A. INTRODUCTION AND PREVIOUS RESULTS

RANGE determination by radar depends on the velocity of electromagnetic waves in air, which is usually obtained from the velocity in vacuum and the dielectric constant of the air under the prevailing meteorological conditions. Aslakson[1] has reversed this procedure by radar ranging over an accurately surveyed path to determine the velocity in air, and using the dielectric constant to reduce his results to a velocity in vacuum. For both these purposes the attainable accuracy is such that errors in the dielectric constant of air are an important contribution to the over-all error, so that more accurate determinations of the dielectric constant of air are needed.

For nonpolar gases the dielectric constant is a function of density and composition only (at constant frequency), while for polar gases it also depends on the temperature. As water vapor is the only polar constituent of the atmosphere, the constant for air is usually split into two additive components, that for dry air and that for water vapor. As the constant for water vapor at room temperature is much greater than for dry air, a small amount of water vapor in supposedly dry air makes a big difference to the measured dielectric contant, and this probably accounts for the large variations in quoted results for "dry" air.

If K is the dielectric constant of dry air, then $(K-1)$ varies directly as the density, so it is sufficient to quote one value, at N.T.P. Table I summarizes the results of previous work.

According to theory the dielectric constant should be independent of frequency from dc to near the first absorption band, which is above 30,000 Mc/sec. Refractive index measurements in the visible and infrared lie on the other side of the absorption bands, and so are not comparable. Neglecting the result of Kerr,[2] which was obtained by a very inaccurate method, the best of the 0–1 Mc/sec results[3,4] and the 9000 Mc/sec results agree very well, as does the one quoted for 24,000 Mc/sec, but the results for 3000 Mc/sec are high. The method used by Phillips[5] was not very accurate, and his probable error is nearly as great as the difference between his value and the other results, so that this difference may not be significant. The result of Hughes and Lavrench[6] was a preliminary result, which was only published because of the apparent support given it by Phillips' work. The present paper reports the final result of this work, during the course of which it was found that the earlier result was too high because of insufficient drying.

* Now at Southwest Research Institute, San Antonio, Texas.
† Now at National Research Council, Ottawa, Canada.
[1] C. I. Aslakson, Nature **164**, 711 (1949).
[2] F. J. Kerr, Proc. Phys. Soc. (London) **55**, 92 (1943).
[3] Watson, Rao, and Ramaswamy, Proc. Roy. Soc. (London) A, 143, 558 (1934).
[4] L. G. Hector and D. L. Woernley, Phys. Rev. **69**, 101 (1946).
[5] W. E. Phillips, Proc. Inst. Radio Engrs. **38**, 786 (1950).
[6] J. V. Hughes and W. Lavrench, Proc. Inst. Radio Engrs. **39**, 839 (1951).

TABLE I.

Frequency	Observer	Ref.	$(K-1)\times10^4$
0–1 Mc/s	Prior to 1946.		5.4–6.0
	Refs. given in	a, b	Ave: 5.8
0	Van Itterbeek *et al.*	c	5.7
1	Watson *et al.*	a	5.76
1	Hector *et al.*	b	5.67
60	Kerr	d	6.5
3000	Phillips	e	5.98
3000	Hughes *et al.*	f	6.06
9000	Crain	g	5.72
9000	Birnbaum *et al.*	h	5.75
24000	Essen *et al.*	i	5.76

[a] See reference 3.
[b] See reference 4.
[c] A. Van Itterbeek and Joz. Spaepen, Physica **10**, 173 (1943).
[d] See reference 2.
[e] See reference 5.
[f] See reference 6.
[g] C. M. Crain, Phys. Rev. **74**, 691 (1948).
[h] See reference 7.
[i] L. Essen and K. D. Froome, Proc. Phys. Soc. (London), Section B **64**, 862 (1951).

B. METHOD OF MEASUREMENT

Birnbaum *et al.*[7] used two resonant cavities to measure the real part of the dielectric constant, the resonant frequency of one cavity remaining unaltered to serve as a fiducial mark while that of the other cavity was varied by admitting dry air to the previously evacuated cavity. The method used here also employed two resonant cavities in a similar manner, but the method of measuring the change of resonant frequency of the variable cavity was different. The method used here should, with suitable precautions, be capable of greater accuracy than that employed by Birnbaum *et al.*; on the other hand, our method does not permit measurement of the loss factor without considerable additional apparatus.

A block diagram of the system is given in Fig. 1. The microwave oscillator 1 (a 707*B* reflex klystron) was frequency modulated from two sources:

(a) A sine wave of about 1 Mc/sec from the rf oscillator. This produced a carrier and sidebands, separated by a frequency interval equal to the modulating frequency.

(b) A sawtooth sweep extracted from the cathode ray oscilloscope (CRO). This swept both carrier and sidebands, in-step, through a range of a few Mc/sec at a repetition rate of 30 counts/sec.

The output from this oscillator was fed through lossy coaxial cables to the two cavities, and the output of the cavities passed through further lossy cables to the crystal mixer. A second klystron, 2, fed by a stabilized power pack but with no other precautions to ensure frequency stability, supplied *c-w* power to the crystal and heterodyned the cavity outputs to a frequency of about 1 Mc/sec. The output of the crystal was amplified in a video amplifier, and applied without rectification to the CRO.

[7] Birnbaum, Kryder, and Lyons, J. Appl. Phys. **22**, 95 (1951).

With the rf modulation absent, adjustments were such that the response curve of the fiducial cavity was displayed on the CRO, while that of the test cavity was outside the frequency range displayed. The CRO picture was a doubled response curve (above and below the undeflected trace) filled with oscillations, since the signal was not rectified. The rf modulating frequency was then made equal to the difference in resonant frequencies of the cavities, so that at the instant the carrier swept through the resonant frequency of the fiducial cavity, the first sideband swept through the resonant frequency of the test cavity. As the outputs from the two cavities were of different frequencies, the video signals they produced by heterodyning with oscillator 2 differed in frequency, and hence interacted to give beats. The CRO picture was then a doubled response curve filled with oscillations, the envelope being the sum of the amplitudes, with a fainter but quite visible trace inside it marking the envelope of the difference of the amplitudes.

This fainter, inside trace marking the envelope of the difference of the amplitudes was used as a sensitive check of the equality of the R.F. modulation frequency and the difference of cavity resonant frequencies. This method is a modification of that described by Sproull and Linder,[8] and the criterion for correct adjustment is exactly that described by them, i.e., the two-humped "difference curve" is symmetrical if and only if the sweep frequencies (carrier and sideband, respectively) pass simultaneously through the peaks of the response curves of both cavities. Figure 2 gives diagrams showing the appearance of the CRO, in this modified method, for (a) correct and (b) incorrect adjustment. In practice the adjustment was easy to make, and seemed considerably more sensitive than the use of markers (which had also been tried).

In addition to the greater sensitivity of the Sproull-Linder method of comparison, this method of determining dielectric constants has the following advantages:

(a) It is independent of the frequency stability of the klystrons. Variations in their frequencies can cause the

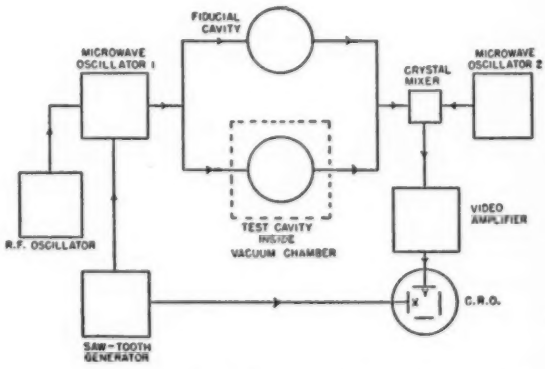


FIG. 1. Block diagram of apparatus.

[8] R. L. Sproull and E. G. Linder, Proc. Inst. Radio Engrs. **34**, 305 (1946).

pattern to wander on the CRO, but does not alter the shape of the pattern, which is the criterion used.

(b) Whereas with frequency markers there are two adjustments (mean frequency and separation), with this system there is only one, the rf oscillator frequency.

(c) The outputs of the two cavities are handled by the same channel, so that differences in delay times are less likely to occur. To test this, measurements were made first with the test cavity, then with the fiducial cavity, having the higher resonant frequency, and no significant difference in the results was found.

The adjustment was made as described above, with the test cavity evacuated, and with it containing air at various pressures up to atmospheric, the frequency of the rf oscillator being read each time.

C. THEORY

Let f_0=resonant frequency of evacuated cavity. f_p=resonant frequency of cavity containing dry air. $\Delta f=f_0-f_p$, P=pressure of dry air in cavity. T=absolute temperature. K=dielectric constant of dry air at P, T. Then,

$$(f_0-f_p)/f_p=K^1-1. \tag{1}$$

Since K is very close to unity, this can be written

$$\Delta f/f_0=\tfrac{1}{2}(K-1). \tag{2}$$

If the gas obeys the ideal gas laws, $(K-1)\propto P/T$. By using P_N, T_N for values at N.T.P., these equations give

$$\Delta f=\frac{f_0}{2}\frac{T_N}{T}\frac{P}{P_N}(K_N-1).$$

Since T remained constant through a run, a plot of Δf against P should give a straight line, from the slope of which line and the observed values of T and f_0, a value of K_N could be obtained.

D. APPARATUS DETAILS

The test cavity was a 10-cm "echo-box," fitted with a second coupling loop and drilled at the ends to increase air circulation. It was mounted in a vacuum chamber, and coaxial leads from the cavity connected to vacuum tight coaxial plugs in the wall of the chamber. A calibrated thermometer hung beside it in the chamber, and the pressure was read on a U-tube mercury manometer. The video amplifier was a war assets apparatus, stated to be flat to 3.5 Mc/sec. The rf oscillator was a commercial signal generator, with a calibration stated to be accurate to ½ percent. Spot checks against WWV showed this accuracy to be attained on the frequencies which could be checked. The fiducial cavity was a calibrated wave meter‡ from which the frequency at which measurements were made could be obtained to 0.01 percent.

‡ Kindly loaned by Admiralty Signal and Radar Establishment, England, at the request of the Defence Research Board, Ottawa.

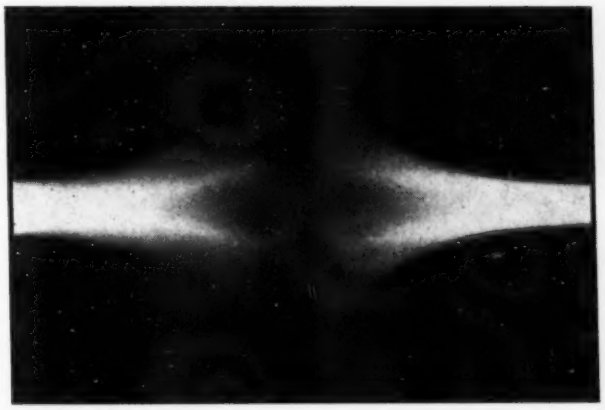


FIG. 2(a). Appearance of the CRO for correct adjustment.

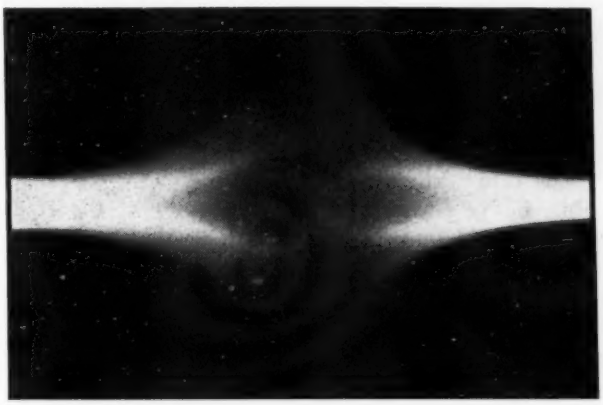


FIG. 2(b). Appearance of the CRO for incorrect adjustment.

Tests for frequency pulling, and for systematic errors, were made by varying the center frequency; initial difference of frequency between cavities; and amplitude of heterodyning signal used. They all failed to reveal any systematic errors.

E. AIR DRYING

Watson *et al.*[5] have pointed out the necessity for extreme care in drying the air, and this was verified here. It was found that the value obtained for $(K-1)$ was the quickest check on the degree of drying obtained, low values corresponding to good drying. In fact, an apparatus of this type, using 3-cm or shorter waves for compactness, could be made into a very sensitive hygrometer. The value of 6.0×10^{-4} for (K_N-1) found by Hughes and Lavrench[6] could be reproduced at any time by using only calcium chloride tubes for drying. Phosphorus pentoxide very quickly became ineffective, presumably because of coating of the active grains by sticky phosphoric acid. Our final system allowed atmospheric air to pass through calcium chloride, then to bubble through two bottles of concentrated sulfuric acid; then to pass through a tube of anhydrous sodium carbonate (to remove any acid spray), and two tubes of phosphorus pentoxide. A capillary tube kept the

linear rate of flow in the drying tubes down to about 5 cm/sec (volume flow 2 l/min). The dried air was kept in a 60-liter reservoir from which it was withdrawn as required. A trap which could be immersed in a bath of solid carbon dioxide in acetone was included near the reservoir, but after several hours operation no frost was visible inside the trap, so its use was discontinued.

F. RESULTS AND DISCUSSION OF ACCURACY

Eight runs were taken, with ten different pressures for each run. The slopes of Δf against P were determined by the method of least squares, and after reduction to 0°C the mean was taken. The probable error of each slope was calculated from internal consistency (fit of the line) and external consistency (agreement of the eight slopes). The ratio of these two estimates differed from unity by an amount which was barely significant.[9] The final result of this analysis, with the probable error of the mean, was

$$(K_N-1)\times 10^4=5.693\pm 0.014.$$

The calibration error of the wave meter was negligible, but a ½ percent calibration error of the rf oscillator could introduce an additional ±0.03. Adding these errors gives the final result: For dry air at N.T.P.

$$(K-1)\times 10^4=5.69\pm 0.04.$$

With a more accurate rf oscillator, this method should give $(K-1)$ to a precision of 0.2 percent. With sufficient care, and more readings, a precision of 0.02 percent (four significant figures) might be attainable, but for this the following points need attention:

(1) The results should be reduced by using van der Waals' equation instead of the ideal gas laws.

(2) Equation (1) is correctly written $(f_0-f_p)/f_p=(K_eK_m)^{\frac{1}{2}}-1$, where K_e is the dielectric constant and K_m the relative permeability. K_m can be put equal to unity for three-figure accuracy in (K_e-1), but not for four-figure accuracy.

(3) The approximation involved in going from Eq. (1) to Eq. (2) is not valid when four figure accuracy is required.

(4) The resonant frequency of a cavity varies with temperature because of thermal expansion of the metal. Using two similar cavities, this effect cancels provided no temperature difference between the cavities occurs. To achieve four figure accuracy in $(K-1)$, this temperature difference must be less than 0.01°C.

(5) The dimensions of the test cavity change with pressure. Enclosure of the cavity in a separate vacuum chamber eliminates the "aneroid barometer" effect, but the space inside the cavity still experiences the same change as would a block of the wall metal, of the same size, under the applied hydrostatic pressure change.

(6) The fiducial cavity should not be open to the atmosphere, but should be evacuated or filled with some nonpolar gas under fixed conditions.

G. ACKNOWLEDGMENTS

The authors wish to thank the Defence Research Board of Canada for a grant to J. V. Hughes for the purchase of apparatus, and for permission to publish this report; and the Research Council of Ontario for a scholarship which enabled H. L. Armstrong to undertake this investigation.

[9] R. T. Birge, Phys. Rev. **40**, 207, 1932.

ASME Symposium on Shock and Vibration Instrumentation Planned for 1952

THE Applied Mechanics Division of the American Society of Mechanical Engineers announces plans for a Shock and Vibration Instrumentation Symposium to be held during its summer meeting at Pennsylvania State College, State College, Pennsylvania, June 19–21, 1952.

Although the principles of operation, sources of error, and methods of calibration of instruments will be stressed, a well-rounded program will be presented covering all fields of interest relating to current shock and vibration testing.

The program includes papers in the following categories: (1) Survey and historical; (2) One-degree-of-freedom instruments; (3) Sources of error and secondary effects; (4) Peak reading gauges; (5) Calibration; (6) Applications. The symposium will close with a general discussion intended to bring out the difficult and intriguing problems which are in urgent need of solution in the instrumentation field and thus to stimulate further development.

It is expected that the symposium papers and discussion will be published in a single volume which will present the current state of knowledge on shock and vibration instrumentation.

The Computation of Dielectric Constants

R. M. REDHEFFER*
University of California, Los Angeles, California

AND

R. C. WILDMAN† AND V. O'GORMAN
United Shoe Machinery Corporation, Beverly, Massachusetts

(Received June 23, 1951)

We consider the shorted-line method as applied to general samples, not restricted to quarter-wave, half-wave, thin, or low loss. The complex dielectric constant can be obtained by combination of two measurements, the sample being followed by short and open circuit, respectively. The equation is simple, non-transcendental, and unambiguous (Eqs. (5), (11), (19), (22)). It lends itself to complete graphical representation (Table I, Fig. 2). When only a single measurement is made, as in conventional practice, a change of variable leads to great simplification of detail ((14), (15), (16), Figs. 2, 3, 4). With due regard to the nature and magnitude of experimental errors, we analyze the conditions for optimum graphical representation, and give a reasoned determination of scale ranges. Large scale curves have been constructed, but only reductions are given here.

INTRODUCTION

IN the shorted-line method for measuring dielectric constants one places the sample at a known distance from the short-circuited end of the guide. The width at twice minimum and the minimum position of the resulting standing-wave pattern are then determined. Alternatively, one places the sample in a cavity, observing the resonant length and the width of the resonance curve at half-maximum. Our language here is adapted to the first situation, although the equations apply to both.

The equivalence of these and other methods has been established elsewhere,[1] and discussion of equipment, measurement procedure, wall losses, and the like is also available.[1,2] To avoid duplication we suppose that the reader is familiar with the details, and reproduce only so much as seems desirable for the purposes at hand.

In view of its all-around excellence the shorted-line method is perhaps the most widely used of the many methods extant, the measurement being rapid, simple, and accurate. Its only shortcoming seems to be difficulty of calculation which, at any rate for samples that are unrestricted, often exceeds the difficulty of measurement. Combined with the widespread use of the method, this circumstance makes it worth while to re-examine computation procedure, and provides the motivation for the present work. Since known methods for quarter-wave samples, low loss samples, and other special cases are satisfactory, we confine our attention to methods that are general. Also, a graph or table is acceptable only if it is valid for all types of guide.

AN ALGEBRAIC OPERATION

The transformation from reflection to impedance occurs so often in this work that we introduce a special notation for it, defining the symbol $[\![\ ]\!]$ by

$$[\![a]\!]=(1-a)/(1+a) \tag{1}$$

for any complex number $a\neq -1$. The utility of this notation depends on the following algebraic properties:

$$\begin{aligned} [\![-a]\!]&=1/[\![a]\!] \\ [\![[\![a]\!]]\!]&=a \\ [\![1/a]\!]&=[\![-a]\!] \\ [\![(a+b)/(1+ab)]\!]&=[\![a]\!][\![b]\!] \\ [\![e^{2a}]\!]&=-\tanh a. \end{aligned} \tag{2}$$

The relation of dielectric constant to interface reflection is $r=[\![\epsilon^{\frac{1}{2}}]\!]$ for free space and

$$r=[\![[(\epsilon-p)/(1-p)]^{\frac{1}{2}}]\!] \tag{3}$$

in guide. The relation of reflection to standing-wave ratio is

$$1/(SWR)=E_{\min}/E_{\max}=[\![|R|]\!]. \tag{4}$$

In (3) the constant p is the squared ratio of free-space to cut-off wavelength,

$$p=(\lambda/\lambda_c)^2. \tag{5}$$

THEORY

By inspection of Fig. 1 we find the standard results.[1,3]

$$R=r+\frac{(1-r^2)\rho e^{2ikd}}{1+r\rho e^{2ikd}}=\frac{r+\rho e^{2ikd}}{1+r\rho e^{2ikd}} \tag{6}$$

upon recalling that the interface transmission $t=1\pm r$. The isolated term r arises from the initial reflection at

* Redheffer's part in this research was also carried out at the United Shoe Machinery Corporation.

† Now at Consolidated Vultee Aircraft Corporation, San Diego, California.

[1] L. G. Montgomery, *Technique of Microwave Measurements* (McGraw-Hill Book Company, Inc., New York), first edition, Chapter X. This reference contains a bibliography of over one hundred titles.

[2] S. Roberts and A. von Hippel, J. Appl. Phys. 17, 610 (1946).

[3] J. A. Stratton, *Electromagnetic Theory* (McGraw-Hill Book Company, Inc., New York, 1941).

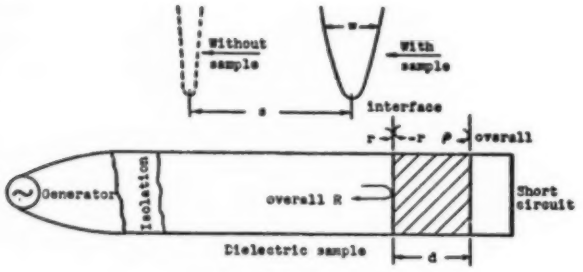


FIG. 1. Notation.

the first interface; the denominator takes account of multiple reflection occurring within the sample. Of course, the expression can also be obtained by setting up the boundary conditions,[2,3] although the derivation is longer. In our notation, (6) becomes simply

$$[R]=[r][\rho\exp(2ikd)] \tag{7}$$

by virtue of (2), and this will be taken as fundamental henceforward.

The value ρ depends on distance from the sample to the end of the guide (Fig. 1). Since this distance is most conveniently determined by use of a dielectric table in back of the sample, it is best to adhere to a few definite values. Following the literature, we single out the cases of *short circuit* (no table) and *open circuit* (quarter-wave table) as of especial interest; the measured reflection for the first will be denoted by R_s, for the second by R_0. Since $\rho=-1$ or $+1$, respectively, one may write

$$\begin{aligned}[R_s]&=[r][-\exp(2ikd)]\\ [R_0]&=[r][\exp(2ikd)].\end{aligned} \tag{8}$$

UNAMBIGUOUS FORM

If we multiply the two relations (8) there results $[R_0][R_s]=[r]^2$ by virtue of (2), and (3) with $[[a]]=a$ then leads to

$$\epsilon=(1-p)[R_s][R_0]+p. \tag{9}$$

Writing $i[R]$ in polar form:

$$i[R]=Pe^{iQ} \tag{10}$$

and separating gives

$$\begin{aligned}\epsilon'&=-(1-p)P_sP_0\cos(Q_0+Q_s)+p\\ \epsilon''&=-(1-p)P_sP_0\sin(Q_0+Q_s),\end{aligned} \tag{11}$$

where

$$\epsilon=\epsilon'+i\epsilon''. \tag{12}$$

We shall see later that P and Q are easily determined from the data (Fig. 2). The choice of variables (9) allows us to assume $-\pi/2<Q<\pi/2$, so that range of the graphs is extended by writing $-Q^*$ for Q^* as in Table I. Complete representation is obtained if we permit positive as well as negative P, and computation of expressions like $\pi/2+Q$ or $\pi-Q$ is thus obviated. The inherent ambiguity of polar representation has been exploited in this way throughout the discussion. It may be noted that the conditions $\epsilon'>1$, $\epsilon''>0$ give the proper sign in (10) by inspection.

TABLE I. Transformations for efficient graphical representation of $i[R]=Pe^{iQ}$.

Range of $8x$	0 - 1	1 - 2	2 - 3	3 - 4	
$x^*=$	x	$1/4-x$	$x-1/4$	$1/2-x$	
$P_0=$	P^*	$1/P^*$	$-1/P^*$	$-P^*$	Periodic
$Q_0=$	Q^*	Q^*	$-Q^*$	$-Q^*$	
$P_s=$	$-1/P^*$	$-P^*$	P^*	$1/P^*$	
$Q_s=$	$-Q^*$	$-Q^*$	Q^*	Q^*	

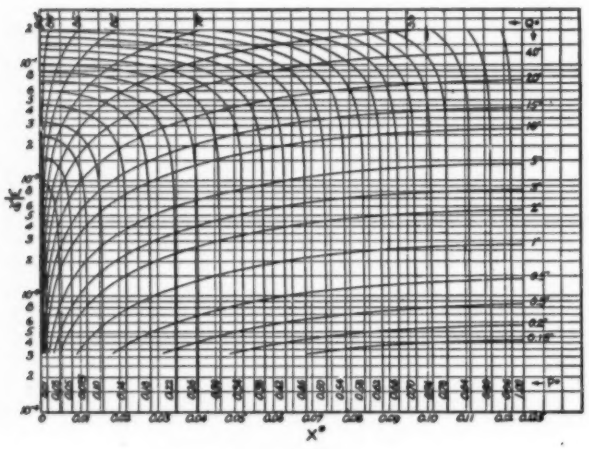


FIG. 2. Chart of the function,

$$P^*e^{iQ^*}=\frac{\tan 2\pi x^*+i\left(1+\csc^2\frac{\pi w}{\lambda_g}\right)^{-\frac{1}{2}}}{1-i\left(1+\csc^2\frac{\pi w}{\lambda_g}\right)^{-\frac{1}{2}}\tan 2\pi x^*}.$$

These relations are not transcendental, they lend themselves to graphical methods, and they are unambiguous; that is to say, Eq. (11) involves single-valued functions only. There is no question of picking the correct Riemann surfaces, and, hence, one need not know even approximately the range of values in which ϵ should lie. Also, in so far as we can determine, the errors have a satisfactory behavior. The equations tend to emphasize the short or open circuit method, according as the one or the other would be more accurate for the conventional procedure (i.e., when used alone). We remark that Eq. (11) does not involve the sample thickness d.

SINGLE MEASUREMENT

The foregoing advantages are compensated to some extent in that the sample must be measured twice; in principle a single measurement ought to suffice whenever ϵ is approximately known. By Eqs. (2) and (8) we have

$$1/[R_s]=-(\tanh ikd)/[r], \tag{13}$$

where the reciprocal is introduced to make the equa-

tions fall into the standard pattern.[2] After using $k=2\pi(\epsilon-p)^{\frac{1}{2}}/\lambda$ and Eq. (3) this becomes

$$\frac{\tan 2\pi(C+iD)^{\frac{1}{2}}}{2\pi(C+iD)^{\frac{1}{2}}}=-\left(\frac{\lambda_g}{2\pi d}\right)\frac{e^{-iQ_s}}{P_s}, \tag{14}$$

where λ_g is the guide wavelength, $\lambda_g=\lambda/(1-p)^{\frac{1}{2}}$, and

$$\begin{aligned}\epsilon'&=(\lambda/d)^2C+p\\ \epsilon''&=(\lambda/d)^2D.\end{aligned} \tag{15}$$

The same procedure with R_0 gives

$$\frac{\cot 2\pi(C+iD)^{\frac{1}{2}}}{2\pi(C+iD)^{\frac{1}{2}}}=\left(\frac{\lambda_g}{2\pi d}\right)\frac{e^{-iQ_0}}{P_0}. \tag{16}$$

The two equations are plotted in Figs. 3 and 4, respectively. It is by the nature of things impossible to achieve complete representation with this approach to the problem, and we have simply selected the first quarter-wave range in which the methods are accurate.[2]

As to content, Eq. (14) is of course equivalent to those of the literature; but as to form, introduction of the square root leads to a notable simplification in practice. Once C and D are determined, ϵ' and ϵ'' are given easily by Eq. (15); this is not the case when the left of Eq. (14) is written $(\tanh ue^{iv})/(ue^{iv})$. Since Eq. (14) is transcendental a graph must be prepared anyway, and the square root may as well be built into the curves from the beginning. A substantial part of the subsequent calculation is thus eliminated.

NONUNIFORM SAMPLES

Sometimes we wish to study a sample which varies in composition as one proceeds along the wave guide axis. Such a sample might be an expanded material, a piece of sandwich radome, or an absorbing surface. How shall one obtain the dielectric constant at point x within the material?

Elsewhere‡ we have shown that

$$d[\![R]\!]/dx=(2\pi i/\lambda_g)[1-[\![R]\!]^2(\epsilon-p)/(1-p)], \tag{17}$$

where x is the sample thickness. Solving for ϵ gives

$$\epsilon=\overline{(-R)^2}(1-p)(1+i\lambda_g\bar{R}'/2\pi)+p, \tag{18}$$

and hence a curve of $\bar{R}$ *versus* x will suffice. The differentiation can be carried out graphically if $\bar{R}$ is presented in Cartesian coordinates, and transferring to polar coordinates then simplifies the multiplications. It must be admitted that such a method is laborious, and limited in accuracy by $\bar{R}'$; but it does give a means of finding information that apparently can be had in no other way. The only approximation in (18) is neglect of wall losses, which can be taken into account by standard methods.

‡ Submitted to the Inst. of Radio Engrs. (abstract).

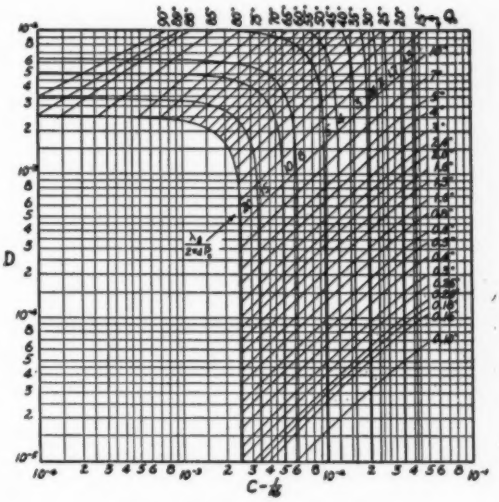


FIG. 3. Chart of the function, $\dfrac{\tan 2\pi(C+iD)^{\frac{1}{2}}}{2\pi(C+iD)^{\frac{1}{2}}}=-\left(\dfrac{\lambda_g}{2\pi d}\right)\dfrac{e^{-iQ_s}}{P_s}$.

THE DATA

The primary measured quantities are the free-space wavelength λ, the guide wavelength λ_g, the cut-off wavelength λ_c, the sample thickness d, the width at twice minimum w, and the shift in minimum position s. The value of w must be corrected for wall losses; the shift s is found on the understanding that the minimum moves toward the short-circuited end upon insertion of the sample (Fig. 1). Our purpose now is to find $i[\![R]\!]=Pe^{iQ}$ in terms of these measured quantities. One may write[1] by Eqs. (4) and (2),

$$[\![|R|]\!]=(1+\csc^2\pi w/\lambda_g)^{-\frac{1}{2}}, \tag{19}$$

which $\cong\pi w/\lambda_g$ for the (usual) case $w\ll\lambda_g$. Also

$$\begin{aligned}\angle R_0&=4\pi(s_0+d)/\lambda_g\\ \angle R_s&=4\pi(s_s+d)/\lambda_g+\pi,\end{aligned} \tag{20}$$

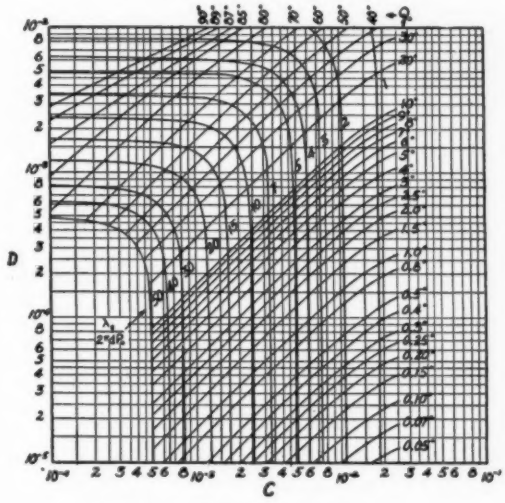


FIG. 4. Chart of the function, $\dfrac{\cot 2\pi(C+iD)^{\frac{1}{2}}}{2\pi(C+iD)^{\frac{1}{2}}}=\left(\dfrac{\lambda_g}{2\pi d}\right)\dfrac{e^{-iQ_0}}{P_0}$.

as is seen by noting $\angle R = 4\pi s/\lambda_g + \text{constant}$, and determining the constant by using a sample with $\epsilon = 1$. From

$$R = [[|R|]][-i\tan(\angle R/2)]$$

one obtains[2] by (2):

$$[R] = \frac{[|R|] - i\tan(\angle R/2)}{1 - i[|R|]\tan(\angle R/2)}. \tag{21}$$

In view of Eqs. (19) to (21), it is easy to plot P and Q *versus* w/λ_g and $(s+d)/\lambda_g$ on a single figure; the same curves suffice for both open and short circuit. Figure 2 is such a graph; it shows

$$P^* e^{iQ^*} = \frac{\tan 2\pi x^* + i[|R|]}{1 - i[|R|]\tan 2\pi x^*} \tag{22}$$

with $[|R|]$ given by Eq. (19). The relation to (20), (21) is given by Table I, in which we have introduced the abbreviation,

$$x = (s+d)/\lambda_g. \tag{23}$$

DETAILS OF GRAPHICAL PRESENTATION

The accuracy with which w can be measured is approximately proportional to the slope, and this in turn is proportional to w. Hence, w should be plotted logarithmically. For x, on the other hand, the behavior of low loss samples shows that a linear scale must be used. As far as the range is concerned, all experimental situations are met by $0.0001 \leqslant w/\lambda_g \leqslant 0.1$ and $0 \leqslant x \leqslant 0.125$. The sufficiency of this last follows from the fact that replacing x by $\frac{1}{4} - x$, $x - \frac{1}{4}$, $\frac{1}{2} - x$, $x - \frac{1}{2}$, and so on, merely changes P into $\pm P$ or $\pm 1/P$, and Q into $\pm Q$, according to a definite pattern (Table I). This symmetry of the transformation allows us to account for every possible situation by a graph of reasonable size. Such a graph has been constructed, consisting of 10 sheets each 16 by 11 inches; the whole curve is shown (to a different scale) in Fig. 2. At all parts of the large curves the accuracy is commensurate with that inherent in the data. Incidentally, the need for accuracy prevents our using polar form and the fact that $[R]$ is a circle for $|R|$ constant.

When only a single measurement is made, then the curves must be supplemented by the two forms (14), (16). Unfortunately one cannot combine these into a single figure, nor is complete representation possible. One can decide on the optimum scale and range, however; it turns out that both C and D should be plotted logarithmically, the former, which is associated chiefly with ϵ', requiring a larger scale. The range is known in so far as one knows the values of ϵ', ϵ'', and d that one is likely to meet; it is found that $10^{-5} \leqslant C \leqslant 1.5$ and $10^{-8} \leqslant D \leqslant 15$ covers most situations. Curves showing one branch for part of this range are given in Figs. 3 and 4; these too have been enlarged for greater accuracy. We use $C - \frac{1}{16}$ as variable in Fig. 3, rather than C. The object is to move the singularity to the origin of coordinates; for a logarithmic plot this makes a great difference in the appearance of the curves and in the accuracy with which they can be read.

SPECIAL CASES

The equations simplify for certain special cases, and a number of particular relations has been derived. It does not enter our plan to pursue the subject here, but we give two examples that seem interesting.

When the loss is low then $|R| \cong 1$, and Eqs. (2) and (9) give

$$(\epsilon' - p)\tan 2\pi x_s = (1-p)\tan 2\pi x_0, \tag{24}$$

where x_s and x_0 refer to open and short circuit. This gives an immediate evaluation of ϵ' in terms of measured quantities.

Again, when $p = 0$, as is the case in coaxial line, then Eq. (11) gives the rather curious result,

$$\tan\delta = \tan(Q_0 + Q_s), \tag{25}$$

for the loss tangent ϵ''/ϵ'. Similar simplifications have been found for Eqs. (14) and (16), but they are probably well known to workers in this field.

ACKNOWLEDGMENTS

The authors are indebted to Stella B. Smith for the lengthy computation and plotting of points for the charts, and to Harold C. Maxfield for measurements of numerous samples.

Contribution to the Theory of the Silicon Carbide Contact

ELSE HOLM
Stackpole Carbon Company, St. Marys, Pennsylvania
(Received November 12, 1951)

1. The high limiting contact voltage of 40 to 50 volts in contacts between silicon carbide members (single crystals of the black-green variety) is explained on the basis of the known temperature-voltage relation.
2. Practically clean silicon carbide possesses a superficial layer which functions as a potential hill for the electrons. Because of its thinness the hill is penetrable for the electric current by means of the tunnel effect. The tunnel resistance per unit area was investigated with terminal areas of the order of 2 mm². Being constant at small voltages, it drops down through some orders of magnitude within the voltage range of 0.5 to 5 volts, showing that the work function for emission of electrons into the barrier material from the adjacent metal is of the order of 3 electron volts. From the amount of the tunnel resistance it is concluded that the thickness of the potential hill is of the order of 10 to 20 Angstrom units. The simultaneity of applied voltage and tunnel current is demonstrated.

GENERAL INTRODUCTION

THE property of ceramic bonded grains of silicon carbide materials to exhibit an electric resistance which decreases with rising applied voltage has long been used for voltage limitation, for example in lightning arresters. It was understood that the limitation effect was caused by the negative temperature coefficient of the resistance of SiC, but there has not been given a quantitative explanation for the limiting voltage per contact. However, some years ago Ragnar Holm suggested that the contact voltage across SiC contacts may be a consequence of the so-called $U-\theta$ relation,[1] which is a relation between the supertemperature θ of the contact surface above the bulk temperature of the contact members and the voltage U across the contact constriction.[1]

In a contact between SiC grains, a small volume around the contact area, the "constriction region," which exhibits the main resistance of the contact arrangement, will play the essential part for the limitation of U (see Fig. 1, where a contact constriction of a symmetric contact is sketched). When the voltage U reaches about 45 volts the constriction gets very hot in the surrounding of the contact surface, and the constriction resistance decreases rapidly at further increase of the current, thus limiting the voltage. It is the aim of the first part of the present investigation to explain the existence of the high limiting voltage U (about 45 volts per contact) with the aid of the $U-\theta$ relation.

A second striking problem lies in the indication that another phenomenon plays a part at relatively low voltages. Some investigators have made the assumption that a part of the SiC contact resistance is the result of the tunnel effect through a thin superficial layer with a very low conductivity of its own. However, there has been little quantitative knowledge about this subject. The second part of this paper is devoted to completing investigations on this low voltage effect.

[1] Ragnar Holm, *Electric Contacts* (Hugo Geber, Stockholm, Sweden, 1946), pp. 31 and 187. As for the definition of the "constriction," see the same book, paragraphs 1 and 3.

I. EXPLANATION OF THE LIMITATION OF THE CONTACT VOLTAGE BETWEEN SiC MEMBERS

A. Introduction with Fundamental Formulas

The $U-\theta$ relation of a symmetric contact between isotropic members (see Fig. 1) is expressed by Eq. (1).

$$U^2=8\int_{T_0}^{T_0+\theta}\rho\lambda dT. \qquad (1)$$

T_0 is the temperature in two equipotential surfaces ("end surfaces" of the constriction), one on either side of the contact, where the electrical gradient may be regarded as negligibly small. Here also the temperature gradient will be negligible and T_0 may be regarded as the temperature of the main mass of the contact members. Between these surfaces of the constriction the voltage U is measured (see Figs. 1 and 4); ρ is the electric resistivity and λ is the thermal conductivity of the contact material within the constriction region, both being functions of the temperature. Since a constriction

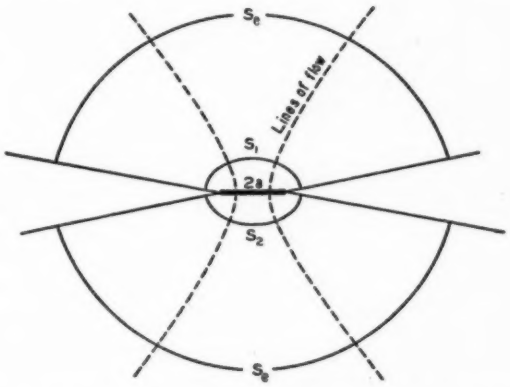


FIG. 1. Schematic design of an electric contact constriction. The heavy line $2a$ is the diameter of the contact surface. (S_e, S_e) represent the so-called end surfaces of the constriction where the electric and the thermal gradient are regarded as negligible. There the temperature is approximately the bulk temperature of the contact members. The contact voltage U is measured between the equipotential surfaces (S_e, S_e). S_1 and S_2 represent the equipotential surfaces between which half of the constriction resistance is localized.

FIG. 2. Schematic design of a long, rectangular sample with zonal silver coatings; areas about 0.02 cm².

may have a very small volume (dimensions as comparable with those of the small contact surface), the condition of isotropy within this volume may be fulfilled.

Certainly, the relation shown in Eq. (1) will lose its validity if the voltage drop caused by a film[2] in the contact is comparable with the constriction voltage. We, indeed, shall find that "films" are present in the contacts in question but do not play any essential part at such high voltages as we use for the verification of Eq. (1) (see Secs. D and E of Part II).

In order to apply Eq. (1), ρ and λ must be determined as functions of T within the temperature range in question. We have measured ρ up to about 800°K but we have not measured λ. However, λ probably varies little with the temperature. According to the tables of Landolt-Börnstein[3] the value $\lambda=0.2$ w/cm deg for SiC of the black variety will be correct within an uncertainty of 20 percent for the temperature range in question.

With the assumption of a constant thermal conductivity, Eq. (1) can be written

$$U^2=8\lambda\int_{T_0}^{T_0+\theta}\rho dT. \tag{2}$$

If ρ is plotted against T in a Cartesian diagram, the integral in Eq. (2) represents an area. Figure 3 provides an example. For instance, the area between the T-axis and the lower curve in the temperature range $T=300$ to $T=2000$°K has the value of 1080 ohm cm deg giving $U=42$ volts.

The temperature of 2000°K is of special significance. That the temperature $T_0+\theta$ in the contact surface is high at the moment when the voltage limitation takes place is indicated by the fact that such voltages are accompanied by the emission of light[4] of red to white color. With dc, first a red glow is seen around the contact region, which at slightly increasing current, changes into a white-yellow color at the moment of the most rapid resistance drop, which may indicate a contact surface temperature of about 2000°K. Other qualitative measurements[5] show a steep drop of resistance on Globar heating elements when the temperature amounts to about 2000°K. We interpret this as the beginning of intrinsic conductivity in SiC.

Equation (1) refers to the steady state and since our contact measurements have been made with 60-cycles ac the question arises whether the temperature of the constriction can follow the heating effect of the current. We shall find that this is the case when the dimensionless function[6]

$$z=2.47\lambda t/ca^2\geqq 500. \tag{3}$$

Here λ is the thermal conductivity, c the heat capacity of the unit volume, a the radius of the contact surface and t the time which is regarded for the development of the temperature. This condition was fulfilled for all ac measurements used in the following.

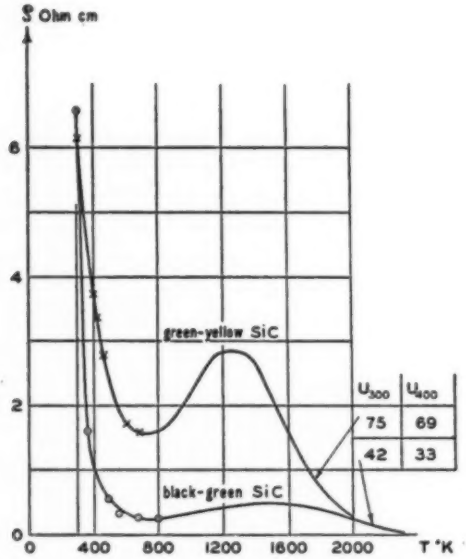


FIG. 3. Resistivity ρ as a function of temperature. Circles and crosses are points measured on SiC crystals of the black-green and green-yellow variety, respectively. U_{300} means the limiting voltage in the case when the bulk temperature T_0 is 300°K.

B. Measurement of ρ in SiC

Method

A current I may flow through a relatively long, rectangular single crystal of SiC and the voltage V may be measured between probes having the distance s from each other. If q is the average cross section of the sample we have $\rho=qV/sI$. Different s values were chosen along the samples.

A difficulty arises in the construction of suitable connections to the sample because of a poorly conducting surface layer on SiC. This was overcome in the following way. A relatively long sample was coated at its ends with a silver paste in such a manner that the silver obviously contacted the whole coated area and ad-

[2] E. Justi and H. Schults, *Neue Versuche zur Deutung der Feinwanderung in elektrischen Abhebekontakten* (Braunschweig. Wiss. Ges. 1, 1949), p. 89, particularly p. 96. The completed expression is due to M. Kohler.

[3] See Landolt-Börnstein, Physik.-Chem. Tabellen, 2390–2391 (5. Auflage).

[4] Best to see with short dc impulses.

[5] Erle I. Shobert II, Stackpole report, Dec. 16, 1949.

[6] Ragnar Holm, *Electric Contacts* (Hugo Geber, Stockholm, Sweden, 1946); see paragraph 11, Eqs. (11.13) and (11.19) and Fig. (11.15), dashed curve.

hered strongly to it. We are able to prove that practically the whole coated area makes electric contact with the silver (see Sec. II).

Nevertheless "film" resistances appeared in the order of 10^4 to 10^5 ohms. But the resistance between a pointed metal probe and the crystal would be much higher,[7] perhaps of the same order as the impedance of the voltmeter used. Therefore zonal silver coatings covering an area of about 10^{-2} cm^2 were employed as probes for our voltage measurements as illustrated in Fig. 2. The resistances resulting from a superficial layer of such areas were of the order of 10^4 ohms which is negligible compared with the 50 or more megohms of the vacuum tube voltmeter. The voltmeter was mounted on an insulating plate and no leakage was observed.

For measurements beyond room temperature, the sample which was prepared in the same manner as described above was mounted on a porcelain plate, and the connection for the current and the voltage were established by means of silver coated springs of stainless steel which remained elastic up to the temperatures that were used. The entire contrivance was put in a quartz tube in an electric furnace. The temperatures were measured by means of a thermo-couple.

Average ρ-values from measurements at different temperatures on black to black-green, and green to green-yellow crystals are plotted in the diagram of Fig. 3. The voltage was measured with an error smaller than 4 percent, and the $V-I$ readings were made at currents small enough not to produce any perceivable heating of the sample during measurement. All ρ-curves which provided the average ρ-values in Fig. 3 are reversible, i.e., measurements with decreasing current give values which lie on the initial curve for rising current. The spread between different measurements on the same crystal is less than 5 percent, and between measurements on different crystals of the same variety less than 10 percent.

C. Contact Measurements

The crystals used in the course of these measurements were of the black to black-green or green to green-yellow variety. Relatively thin pieces of rectangular shape were cut from single crystals so that any eventual film caused by silica was always removed (see Sec. II).

The contact resistance R between two crossed single crystals of SiC (dimensions about $12\times2\times1$ mm^3) was measured (see Fig. 4). Each crystal had silver coatings at its ends for the electric connections and also a silver spot on one side surface. These spots were soldered to a fixed holder and to a balance arm, respectively. It was important to attach the crystals to holders of high heat conductivity to provide good cooling. The possibility existed of bringing different areas of the members into contact. Very likely, any contact consisted of one single

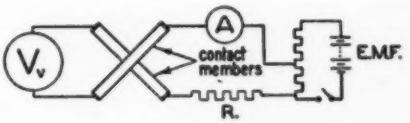
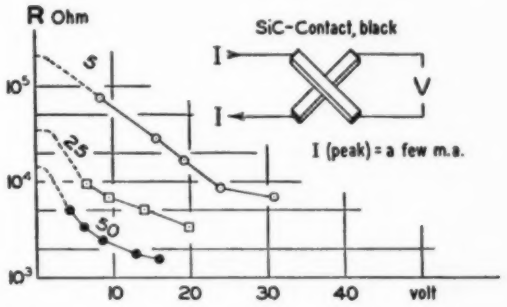


FIG. 4. Wiring diagram for measuring the current I through a contact between two crossed, single crystals of SiC and the voltage U across the contact constriction.

spot because of the existence of only one degree of freedom at contact make. The contact load was produced by a weight. The balance and the devices at which the leads from the balance were attached rested on a rubber plate in order to prevent disturbances of the contact by vibrations. The whole apparatus was electrically screened.

Oscilloscope records of the contact voltage U against the current I were taken with 60-cycles ac. The wiring diagram sketched in Fig. 4 and the schematic design of a contact constriction in Fig. 1 make clear[8] that the $U-I$ records measure merely the so-called contact resistance R. However, R will consist not only of a constriction resistance R_c between the end surfaces S_e of the constriction but of a barrier resistance R_f in the contact as well (see Sec. II). Thus $R=R_f+R_c$. Resistances R calculated from $U-I$ records are plotted



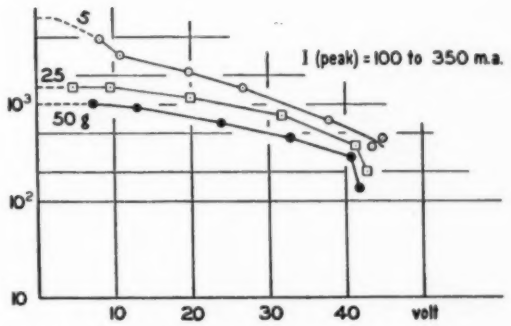


FIG. 5. Contact resistance $R=R_f+R_c$ plotted against voltage U across the contact constriction. Contacts between two crossed, single SiC crystals of the black to black-green variety. Contact load $P=5$, 25 or 50 g, respectively. Upper triplet of curves is taken at small peak currents (a few ma), lower triplet at higher peak currents (about 300 ma).

[7] Jones, Scott, and Sillars, Proc. Phys. Soc. (London) **62**, 333 (particularly 336 and 337) (1949).

[8] Ragnar Holm, *Electric Contacts* (Hugo Geber, Stockholm, Sweden, 1946), Fig. (8.04).

FIG. 6. Oscilloscope $U-I$ record of a SiC contact where U and I are relatively small so that the high temperature in the contact surface, which is the condition for the limitation of U, is not yet reached. I(peak) $=\pm 9.2$ ma. Contact load P about 5 g. $U/I=R_f+R_c$. At voltages up to a few volts, $R_f \gg R_c$. R_f(initial) $>10^6$ ohm.

against contact voltage U in Fig. 5, where each point corresponds to an average of three to four measurements with different contact spots under the same conditions. The contact load was 5, 25, or 50 g. Figures 6 and 7 show typical records, Fig. 6 with a relatively low peak current and low contact voltage (high temperature necessary for limiting U not yet produced in the contact surface) and Fig. 7 with a relatively high peak current and fully established limitation of U.

Discussion of Fig. 5

There is a great uncertainty in the evaluation of the resistance R at low voltages where $R \approx R_f$ (dashed parts of the curves). R is probably constant at voltages up to about 0.5 volt, but the error in determining the initial resistance may sometimes be as high as 100 percent or even more at small contact loads as 5 g.

However, the voltage region which plays the essential part for our contact problem in Sec. I lies beyond 5 to 10 volts, where $R \approx R_c$; and here the spread of R is less than 5 percent. With peak currents larger than 100 ma (see lower triplet of curves) and at, for instance, 7 volts, where very likely $R \approx R_c$, we find R values indicating that the contact surface varies fairly proportional to the load P. The low resistance at these currents proves that the high temperature in the contact has produced a larger conducting contact surface than would correspond to the applied load at small currents (*cf.* R at 7 volts on the upper triplet of curves). According to the low value of the initial resistance at higher currents, this final contact surface remains unchanged when the alternating current decreases during the cycle. Very likely, not only the existent contact surface was increased but also a new one has been added.

However, the feature in which we are mainly interested, in Sec. I, is the marked steep drop which appears between 40 and 50 volts (see lower triplet of curves in Fig. 5, and also Fig. 7). We identify this drop with the resistance drop at the start of intrinsic conduction. The voltage at which it occurs is what we call the "limiting" voltage. The limitation is fairly sharp according to the contact measurements, since they do not include any bulk resistance (see Fig. 4).

It remains to show that the contact temperature followed the current variations with a negligible lag in the next surrounding of the contact surface, say in the surrounding that contains 0.75 of the constriction resistance, which lies within the distance $2.5a$ from the contact surface. We use the same symbols as in Eq. (3). When 90 percent or more of the steady state temperature is reached in less than $t=1/60\times 30=5.6\times 10^{-4}$ sec we regard the lag of the temperature behind the 60 cycle ac as negligible. Now $2.5a$ corresponds to $4b$ in Fig. (11.15) of the book *Electric Contacts*, and the dashed curve of that figure shows that the corresponding z is 500, as is expressed in Eq. (3). According to our Fig. 5 (lower triplet of curves) the contact resistance is $R \approx 400$ ohm at the moment when the U limitation begins. Assuming only one contact surface (most unfavorable case) we find $a=\rho/2R=2.5\times 10^{-4}$ cm, and, with $\lambda=0.2$ and $c=2$, $t=1.3\times 10^{-4}$ sec, which is less than 5.6×10^{-4} sec.

A further increase of the current, peak current 400 to 500 ma or more, usually is followed by decreasing U. This effect we interpret as a consequence of appreciable heating, not only of the constriction, but of the bulk as well, i.e., increasing T_0 in Eq. 2. Actually a high temperature of the contact members was measured. $T_0=400°$K yields $U=33$ volts (see lower curve in Fig. 3).

Application of Eq. (2)

Taking for granted the reference of the steep resistance drop at about 45 volts to the temperature $T_0+\theta=2000°$K in the contact surface we possess the means of applying Eq. (2), if $\int \rho dT$ can be determined. The results of our ρ-measurements which are plotted in Fig. 3 against temperature as circles give average ρ-values up to 800°K. At higher temperatures probe measurements are no longer feasible and we are com-

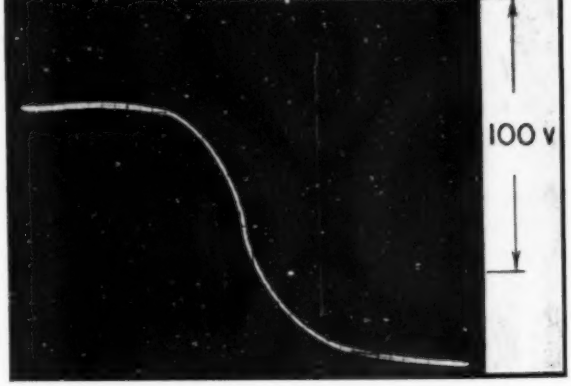


FIG. 7. Oscilloscope $U-I$ record of a SiC contact at such high voltages U that the voltage limitation is fully developed. I(peak) $=\pm 362$ ma. Contact load $P=50$ g. $U/I=R=R_f+R_c$. R(initial) about 1200 ohm. U(limiting) $=47$ volts for forward direction.

pelled to make plausible calculations.[9] This has been done with the following assumptions:

(1) Fully dissociated donors (or acceptors) above 800°K.

(2) Electron mobility proportional to $T^{-1.5}$ and equal to 200 cm²/volt sec at room temperature.

(3) An activation energy for the intrinsic conductivity of 2 ev. Assumption 3 is based on the large drop of ρ at about 2000°K (see end of Introduction to Sec. I).

Assumption 1 is supported by measurements of ρ around 800°K which indicate a minimum of ρ. Here, all donors are supposed to be ionized in the hotter part of the constriction and in the region of the potential hill of the contact. The connection with the measured point at 800°K requires $nb=2\times10^{20}$ at room temperature where n is the number of conducting electrons per cm³ (which is assumed to equal the number of the donors) and b=electron mobility. In order to further agreement with the measured ρ-curve we had to choose[10] $b=200$ cm²/volt sec and $n=10^{18}$ cm^{-3}.

With the aid of the assumptions above we were able to calculate the $\rho-T$ curves beyond the measured points up to higher temperatures, T. Planimetring of Fig. 3, lower curve, yields for the contact temperature $T_0+\theta=2000$°K and $T_0=300$°K.

$$\int_{300}^{2000}\rho dT=1080,$$

and for the limiting voltage $U=42$ volts according to Eq. (2) with $\lambda=0.2$.

This agrees reasonably with Fig. 5, which determines the limiting voltage around 40 to 45 volts.

The error limits of the measurements are, of course, wide. It is not even certain that the contact constrictions in question have quite the same ρ as was determined along the crystals used for the resistivity measurements. But it is obvious that Eq. (1) determines the right order of magnitude of U and thus makes understandable the high limiting contact voltages at silicon carbide contacts.

According to the upper curve of Fig. 3, where average values from our ρ-measurements on green or green-yellow crystals up to 800°K are used in the analysis, the voltage limit should be about 75 volts. This is actually the order we usually measured on such crystals.

II. THE CHARACTER OF THE SURFACE LAYER ON SiC CRYSTALS

A. Introduction

A visible film (thickness 100 to 1000A) is normal for SiC crystals. The removal of this film by means of abrasion in air or by etching does not mean that the surface layer becomes of the same character as the bulk material of SiC. Electric measurements still reveal a surface film or surface layer with a great electric resistance at small voltages. Section II of this paper is devoted to an investigation of the electrical properties of this layer when it is brought into contact with a metal. The resistance at low voltages turns out to be explainable as a "tunnel" resistance across a very thin barrier (thickness of the order of 10 to 20A).

We were not able to make those investigations by means of a contact between single SiC crystals because of the difficulty and possible errors of a direct determination of the contact surface. We therefore investigated single crystals of a rectangular shape with determinable silver coated areas after having proved that an essential part of these areas make contact with the silver. Thus, all resistance measurements in Sec. II refer to contacts between SiC and silver, with a contact surface of about 2×10^{-2} cm².

The problem of the surface layer has to some extent been investigated by Braun and Busch[11] and recently by Jones, Scott, and Sillars.[12] Braun and Busch assume that freshly cleaned silicon carbide surfaces are covered by a thin silica film which produces a tunnel resistance. Jones, Scott, and Sillars compare contacts between a grammophone needle and SiC surfaces of different, determinable cleanliness. With the cleanest surfaces an electron diffraction camera showed sharp silicon carbide patterns with Kikuchi lines. Even an alien film of the thickness of 3A would cause obscuration of these lines. Other surfaces used were covered by evaporation with determinable thickness of silica. The authors report similar electrical properties for the contact between the needle and any of the differently prepared surfaces. The resistance drops through a couple of orders of magnitude when the voltage across the contact increases some few volts. This singular result seems to indicate that the superimposed film was always either mechanically damaged upon making contact or electrically broken down, and that a depletion layer remained active. Unfortunately the authors did not apply the usual precaution against the break down, namely, the use of small electromotive forces.[13] They only limited the current. Because of this omission the results have a limited validity. Nevertheless, they may serve as a warning against too detailed assumptions about the character of the surface layer.

For the following it does not matter whether the singularity of the surface layer, that will be demonstrated, is caused by an alien film or by a depletion layer at the boundary of the lattice. It is sufficient for the treatment that the action of the barrier layer con-

[9] For formulas see, for example, H. C. Torrey and C. A. Whitmer, *Crystal Rectifiers* (McGraw-Hill, New York, 1948) sections 3.2 and 3.4.

[10] G. Bush and H. Labhart, [in Helv. Phys. Acta **19**, 463 (1946)] calculated a much lower b value on the basis of Hall effect measurements, considering either electron or hole conduction. Assuming both kinds of conduction to occur simultaneously the b value that we have assumed is reasonable.

[11] A. Braun and G. Busch, Helv. Phys. Acta **15**, 571 (particularly 582) (1942).

[12] Jones, Scott, and Sillars Proc. Phys. Soc. **62**, 333 (1946).

[13] Ragnar Holm, *Electric Contacts* (Hugo Geber, Stockholm, Sweden, 1946), paragraph 25.

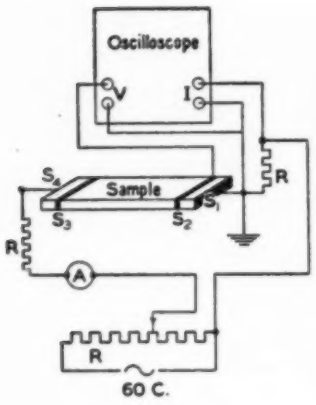


FIG. 8. Wiring diagram for measuring the resistance R_f of the barrier layer in a contact between SiC and metal. S_1 to S_4 are metal coated areas which define the conducting areas between the SiC surfaces and the metal. The voltage V was measured either between S_1 and S_2 or between S_1 and S_3 so that the $V-I$ records determine R_f of the area S_1, in the first case with a short length of bulk, and in the second case with a longer length of bulk.

cerned in a contact can be represented by a potential hill.

In our case where we have a contact between SiC and metal, the potential hill will be much steeper towards the side of the metal (because of the denser charges) than towards the side of SiC, so that the hill in this case will be smaller than in a symmetric contact between SiC crystals.

We make our calculations on the basis that the undisturbed barrier between metal and SiC has an ideal rectangular shape with the thickness of s Angstrom units and the height of φ ev. Its resistance will be designated as the resistance R_f of the barrier. The conversion of the results to another shape of the barrier can easily be made when the size of the rectangular one is known. A greater value of s will then result if φ is supposed to be the same in both cases. For instance, it can be seen from R. Holm's curves[14] that, when taking into account the possibility of rounded off boundaries of the barrier according to the image force, one must make s larger by about 5A in order to obtain the same characteristics as with the rectangular barrier.

Our results give quantitative explanation for a certain barrier layer producing a tunnel resistance. In addition to the older results we are able to provide new and more reliable data as to the thickness of the barrier layer and to the work function for the emission of electrons into the material of the barrier layer from the metallic terminal.

B. Method of Measuring the Resistance of the Barrier Layer in a Contact Between SiC and Metal

The crystals used in the course of the measurements were of the black to black-green variety. The preparation of the samples and the application of metal coatings was the same as described previously for the contact measurements in Sec. I. The crystals were thoroughly cleaned before the metal coating was applied to parts of them. The wiring diagram is sketched in Fig. 8.

Crystals with very smooth surfaces were chosen and it was found that practically the whole silver coated area made electric contact with the silver. The uniformity of the contact with the coating was proved in two ways: (1) Silver coatings on the end surfaces (S_0) of a single crystal were subdivided into smaller areas of about $S_0/2$ and $S_0/4$. The average resistance of those areas changed in the inverse proportion to the area. (2) At about 6 volts on curve I in Fig. 9, when the resistance of the barrier is reduced to a negligible amount and when the sample is heated not more than 10°C beyond room temperature, we calculate $\rho=4$ ohm cm for the portion of the sample between S_1 and S_2 as nearly the same (namely $\rho=300\times2\times10^{-2}/1.4=4.3$) as for the crystal portion between S_1 and S_3 (*cf.* curve III in Fig. 9). These results prove that, under said circumstances, the resistance near S_1 does not comprise a perceivable constriction resistance. In other words, the metal terminal contacts a large part of the coated area.

C. Results of Measurements of the Resistance of the Barrier

The measurements were made with 60-cycles ac by means of oscilloscope records of voltage against current. The resistance was calculated as $R=V/I$ at any time. It consists of a barrier resistance R_f and a bulk resistance R_b, where b refers to bulk, thus $R=R_f+R_b$. Figure 9 summarizes the results of R measurements on

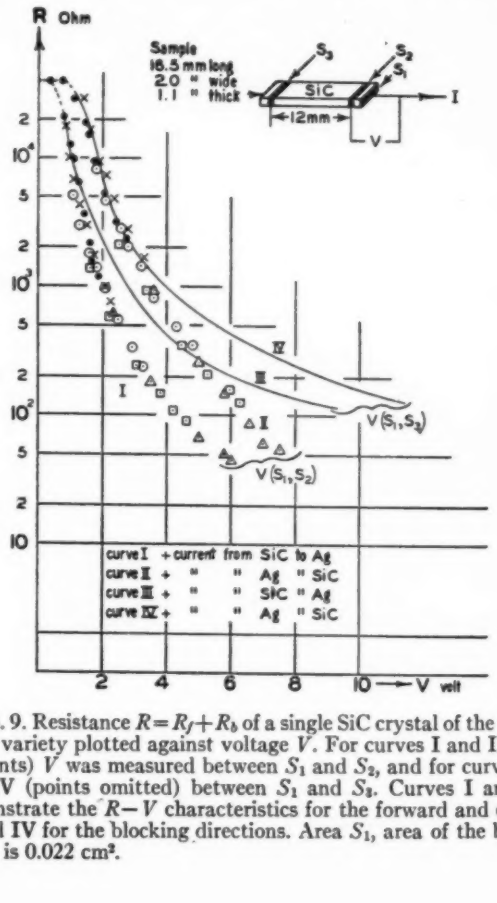


FIG. 9. Resistance $R=R_f+R_b$ of a single SiC crystal of the black-green variety plotted against voltage V. For curves I and II (row of points) V was measured between S_1 and S_2, and for curves III and IV (points omitted) between S_1 and S_3. Curves I and III demonstrate the $R-V$ characteristics for the forward and curves II and IV for the blocking directions. Area S_1, area of the barrier layer, is 0.022 cm².

[14] As for the definition of "tunnel" resistance, see Ragnar Holm, J. Appl. Phys. **22**, 569 (1951).

one crystal of the black to black-green variety. The peak currents lie in the range of 1 ma to 200 ma.

Evidently the records begin at zero current with a finite slope, for example the $V-I$ record in Fig. 10 indicates the average initial resistance to be $\bar{R}_{\text{init}}=10{,}000$ ohms. Other black to black-green crystals possess initial resistances between 10,000 to 40,000 ohms, depending upon the size of the contact area. R_{init} usually stays constant in the voltage range 0.1 to 0.5 volt, which also seemed to be the case with the initial resistance of a contact between SiC pieces (see Figs. 6 and 7). Since the whole end surfaces of the relatively long crystals were covered with silver coatings, these surfaces constitute the electric contact surfaces[15] between metal and SiC. They, of course, remained constant for all currents. Their areas were between 0.01 to 0.045 cm² in all investigated cases.

In Fig. 9, points with different symbols refer to oscilloscope records which differ with respect to the peak

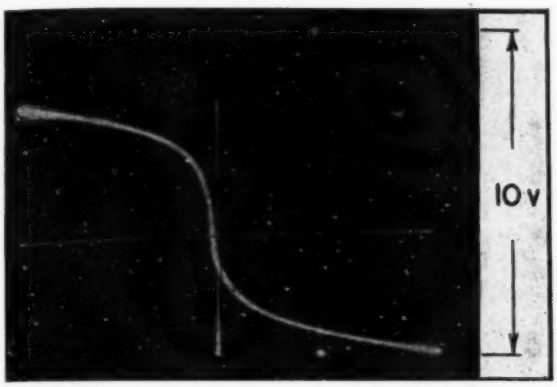


FIG. 10. Oscilloscope $V-I$ record of a single SiC crystal (dimensions: $1.6\times0.2\times0.12$ cm³) of the black-green variety ($\bar{\rho}=6.2$ ohm cm). V is measured across the end surfaces of the crystal. $V/I=R=2R_f+R_b$. I(peak)$=\pm2.83$ ma R(initial) about 20,000 ohm. Tunnel resistance per unit area about $10{,}000\times0.024=240$ ohm cm².

current. The resistance $R=R_f+R_b$ is plotted against voltage V. The coated end areas were 0.022 cm² which we regard as the area of the barrier region. Thus curves I and II represent the resistance of a barrier which extends over an area of 0.022 cm² plus the resistance of about 2-mm length of bulk for the forward and blocking directions, respectively. The only difference between these curves and curves III and IV is that the resistance in III and IV includes the resistance of about 14-mm length of bulk. Comparing curve I with curve III (forward direction) we see that adding bulk resistance to the same barrier resistance makes the resistance drop less pronounced. The reason for this fact obviously is that the added bulk resistance does not vary essentially at voltages up to 10 volts.

[15] See the proof at end of Sec. B.

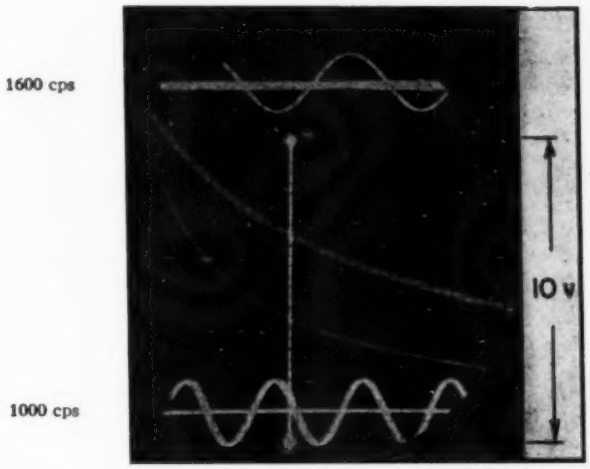


FIG. 11. Single sweeps when a condenser is suddenly discharged through a single crystal of the black-green variety (dimensions: $1.6\times0.2\times0.12$ cm³, and $\bar{\rho}=6.2$ ohm cm). V, the ordinate, is the voltage at the condenser in the moment t. The curves have different time scales. The $V-t$ points (at 10 volts) in the very beginning of the curves, i.e., about 10^{-6} sec after contact closure measure $\bar{R}_b=400$ ohm, the mere bulk resistance. But the $V-t$ point at about 1.8 volt (lower curve) measures essentially the barrier resistances $2R_f$ about $1590-400=1190$ ohm.

D. The Barrier Resistance is a "Tunnel" Resistance

Evidence that the barrier resistance decreases simultaneously with the applied field is given in the following way. A capacitor C of 2 to 4 μF was suddenly discharged through the SiC crystal across S_1 and S_2 (see the sample sketched in Fig. 13). The resistance in such a case is $R=2R_f+R_b$. The discharge is recorded as a single sweep which begins with a time delay of about 10^{-6} second behind the closure of the contact. Two typical records are shown in Figs. 11 and 12.

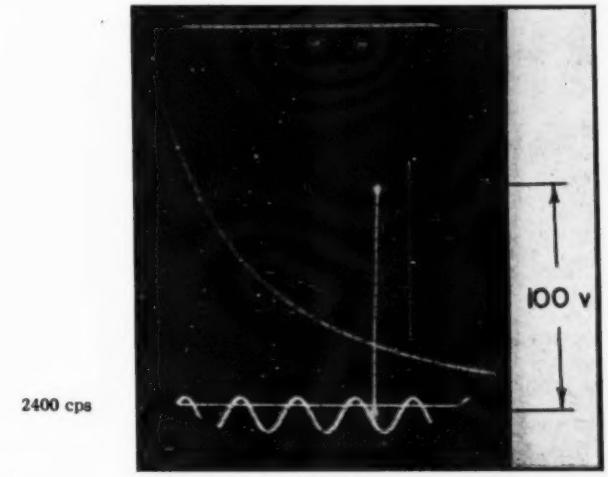


FIG. 12. Single sweep when a condenser is suddenly discharged through a single crystal of green-black variety (dimensions: $1.64\times0.25\times0.18$ cm³ and $\bar{\rho}=12$ ohm cm). All points along the $V-t$ curve in the voltage range of about 140 to 10 volts measure the mere bulk resistance $\bar{R}_b=430$ ohm.

For any moment t the resistance is given by Eq. 4,

$$R=\frac{V}{CdV/dt}, \tag{4}$$

when V is the momentary voltage at the capacitor. dV/dt is the slope of the $V-t$ curve at the point concerned.

Figure 13 presents a typical example of measurements of this kind on a sample (green-black variety), which is sketched in the same figure. The whole resistance $R=2R_f+R_b$ which is calculated according to Eq. (4) is plotted as ordinate against voltage V (solid circles). These points measure, at voltages larger than 10 volts, a constant resistance of about 430 ohms which agrees with the bulk resistance at room temperature of the sample used, with $\bar{\rho}=12$ ohm cm. This fact proves that the discharge through the sample does not perceivably increase its temperature. It is therefore plausible that the resistance drop at voltages below 10 volts must be the result of the decrease of R_f, the barrier resistance.

We are also able to prove that even the first discharge point which corresponds to a current flow for about 10^{-6} sec shows the resistance which corresponds to the voltage at that time. For this reason we have plotted in Fig. 13 results from R measurements[16] on the same sample but with the longer action of dc of about 1 sec (empty circles). At low voltages these points are in good agreement with the solid circles from the discharge case. It is easy to show that in such a large contact surface as 0.045 cm² between SiC and the metal terminal not even a few percent of the supertemperature[17] at dc could be attained during the first part of the discharge; and the said agreement proves that the contact drop concerned was not a result of temperature rise.

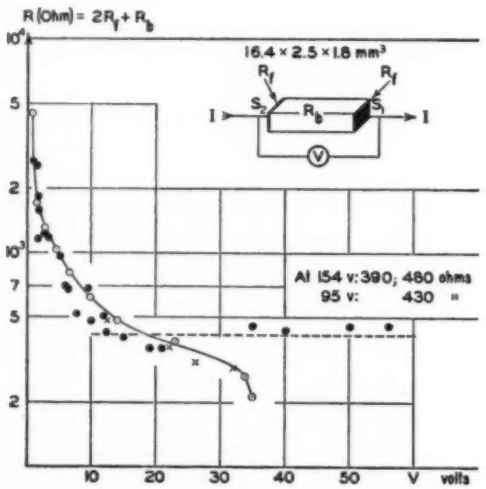


FIG. 13. $R-V$ characteristic, measured on a single crystal of the green-black variety ($\bar{\rho}=12$ ohm cm) by means of sudden discharges of a condenser (see the $V-t$ curve in Fig. 12) through the crystal (solid circles). At such high voltages as 140 to 10 volts all $V-t$ points measure the bulk resistance $R_b=430$ ohm (dashed line). Discharges with smaller voltages yield higher resistances $R=2R_f+R_b$, where now R_b is negligible beside $2R_f$. The empty circles and crosses represent R measurements on the same sample by means of a vacuum tube voltmeter at dc before and after the condenser discharges. The current flow during these measurements lasted about 1 sec.

[16] By means of a vacuum tube voltmeter.

Only the tunnel effect can be the cause of an instantaneous drop of R with voltage application. Since $R=2R_f+R_b$, R_b being constant as far as the bulk temperature T_0 remains constant, the drop will be the result of the decrease of R_f. First, at voltages beyond 5 to 10 volts, R_f is negligible and we measure a constant bulk resistance $R_b\approx 430$ ohms.

The deviation of the dc points from the discharge points at voltages beyond 10 volts is explainable as caused by appreciable I^2R heat which increases the bulk temperature during the time of reading the instrument, which means a decrease of ρ.

E. Determination of the Thickness of the Barrier and the Work Function for the Emission of Electrons into the Barrier Material from the Surrounding Metal

In the preceding sections we have given evidence that the barrier resistance R_f is a tunnel resistance. We have found that the barrier assumes a conductivity corresponding to the applied voltage simultaneously with the voltage application. The barrier resistance seems to be constant at low voltages, then decreases with rising voltage. The resistance R_f is inversely proportional to the area of the barrier region (see end of Sec. B).

In the following we use recently deduced formulas[18] for the tunnel resistance in thin insulator films, whether between metals or semiconductors. Our arrangement with a film between SiC and metal will be intermediate to extreme cases, designated A and B.

(A) Film between semiconducting members. From our experimentally supported assumption for the number n of conducting electrons per cm³ (see Sec. C, Application of Eq. (2)) we find that $n=10^{18}$ at $T=300°$K. Inserting those numbers in R. Holm's equation (30) yields the tunnel resistivity,[19]

$$\sigma=\frac{V}{4.33\times 10^5}\times\exp[0.683(s/V)\{\varphi^{\frac{1}{2}}-(\varphi-V)^{\frac{1}{2}}\}]\text{ohm cm}^2, \tag{5}$$

where V = voltage across the barrier, s = thickness in Angstrom, and φ = work function in electron volts for

[17] At the time t the fraction F of the steady temperature will be reached when

$$1-\exp[-\pi^2\lambda t/4a^2c]\approx F,$$

where a is a linear dimension somewhere between the average radius, 0.1 cm of the silver coating and half the length 0.5 cm of the crystal; say $a=0.2$ cm. We then find that in order to make $F>0.5$, t must be greater than 0.1 sec.

[18] Ragnar Holm, J. Appl. Phys. **22**, 569 (1951).

[19] Corrected by the factor V—*cf.* J. Appl. Phys. **22**, 1217 (1951).

electron emission from the metal terminal into the barrier material.

(B) Film between metal members. We use R. Holm's Eqs. (26) and (27) for j_1 and j_2 with the same designations.

$$j_1=\frac{10^{11}}{3.18s}V(\varphi+V/2)^{\frac{1}{2}}\exp[-1.025s(\varphi+V/2)^{\frac{1}{2}}]$$

$$+\frac{10^{11}}{1.06s^2}\{(\varphi-V/2)\exp[-1.025s(\varphi-V/2)^{\frac{1}{2}}]$$

$$-(\varphi+V/2)\exp[-1.025s(\varphi+V/2)^{\frac{1}{2}}]\}\ \text{amp/cm}^2. \quad (6)$$

$$j_2=1.53\times10^{10}(F^2/\varphi)\exp\left[-\frac{0.68}{F}\varphi^{1.5}\right]\text{amp/cm}^2. \quad (7)$$

The tunnel resistance per unit area is

$$\sigma=V/(j_1+j_2)\ \text{ohm cm}^2. \quad (8)$$

The formulas (5) to (8) show that σ or R_f drops to very small values within the voltage range between 0.5 volt and a little more than the voltage value of the work function φ. The thickness s of the barrier essentially determines the order of magnitude of σ, and φ essentially determines the abscissas for the steepest drop of the σ-curve. Obviously φ is about 3 electron volts and s lies between 10 and 20 A. Figure 14 presents curves which have been calculated on the basis of Eqs. (5) to (8) for different values of the variables s and φ. The solid curves refer to case (A) and the dashed ones to case (B).

The plotted points represent R_f values per unit area from measured $R-V$-curves (see typical $R-V$ curves in Fig. 9) after subtracting the bulk resistance R_b. The resistance drop appears in a voltage range of about 3 volts with the calculated slope; and we conclude that the barrier resistance, being a tunnel resistance, assumes insignificant values at voltages of 4 to 10 volts. This is the reason for referring Eq. (1) merely to the constriction of a contact.

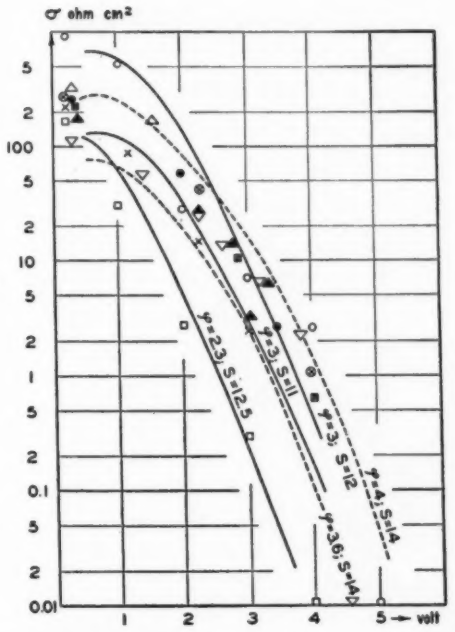


FIG. 14. Calculated tunnel resistance σ per unit area as a function of the voltage V across the potential hill in a contact between two SiC crystals (solid lines) and one between two metal members (dashed lines). The variables are the thickness sA of the hill and the work function φ electron volts. The plotted points represent R_f measurements on different crystals of the black to black-green variety in contact with a metal coating.

ACKNOWLEDGMENT

It is a pleasure to acknowledge the help received from my husband, Dr. Ragnar Holm, through the discussion of these problems. Grateful acknowledgment is also made to Mr. H. M. Dressel and Mr. Erle I. Shobert. The author also wishes to thank the management of the Stackpole Carbon Company for supporting this work and for their permission to publish it.

The Initiation of Electrical Breakdown in Vacuum*

LAWRENCE CRANBERG
University of California, Los Alamos Scientific Laboratory, Los Alamos, New Mexico
(Received October 29, 1951)

The hypothesis is suggested that initiation of high voltage breakdown in vacuum is due to traversal of the high voltage gap by a clump of loosely adhering material. The implication of this hypothesis for uniform-field gaps is that the breakdown voltage is proportional to the square root of the gap length. A summary the literature is presented which supports this conclusion for a range of voltages from 20 kilovolts to 7 megavolts, and for a range of gap distance from 0.2 mm to 6 meters. Additional qualitative evidence is presented which tends to support the proposed hypothesis.

I. INTRODUCTION

THE phenomenon of electrical conduction between metal electrodes in high vacuum has been observed and studied for many years. The conduction is of two kinds: cathodic electron emission, which is very sensitive to the cathodic gradient and occurs only at very high gradient (of the order of megavolts/cm); and a low gradient, heavy current arc in which vapors of the electrode materials play a role. The only method by which the second type of conduction has been obtained has been by application of progressively higher voltage. The onset of the arc limits the practical usefulness of vacuum as electrical insulation and is therefore referred to as breakdown.

In 1928, with the formulation of the Fowler-Nordheim[1] theory of cold electron-emission from metals, a theoretical basis was provided for the interpretation of the first phenomenon. The comprehensive analysis of Stern, Gosling, and Fowler[2] showed substantial agreement of theory and experiment, and the Fowler-Nordheim theory has been regarded as reasonably well established, at least in the domain of low voltage and high gradient.

The initiation of breakdown, which is the concern of this paper, is a matter of much less theoretical interest, but of great practical importance for the design of high voltage vacuum apparatus; and about this there is much less agreement. Up to 1936, with the work of Hull and Burger,[3] Snoddy,[4] Beams,[5] and Ahearn,[6] it appeared to be fairly well established that breakdown was initiated by cold electron-emission from points on the cathode, these electrons producing intense localized heating of the anode. The conditions studied by these workers were those of uniform and cylindrical fields, with high cathode gradient, but low (less than 30 kv) total voltage. Even in this work, however, the evidence has not been entirely concordant: *viz.*, Ahearn's observation of the effect on breakdown of the glass walls surrounding the electrodes, and Beam's calculation that the measured breakdown gradient from a pool of mercury, fitted into the Fowler-Nordheim relation, gives less than 1 electron/cm²/sec as the initiating current density.

In 1935, with the work of Anderson,[7] who used the segmented-tube technique suggested by Van de Graaff to obtain high voltage in vacuum, it became clear that the explanation in terms of cathode gradient, though possibly applicable to some cases, could not explain satisfactorily the results obtained with high voltage gaps. His work, together with data taken subsequently with the same or similar apparatus by Trump and Van de Graaff,[8] showed breakdown to be associated with progressively lower cathode gradient as the gap distance and voltage were increased, and this has been denoted the "total-voltage" effect. Because of the great sensitivity of emitted current to cathode gradient in the Fowler-Nordheim relation, these results cannot be reconciled with an initiation hypothesis which depends on heating of the anode to some critical temperature by field-emitted electrons.

This difficulty has led Anderson and Trump and Van de Graaff to postulate that positive ions as well as electrons play a role, one particle producing the other at opposite electrodes, breakdown resulting when the product of positive-ion yield per bombarding electron at the anode, and electron-yield per positive ion at the cathode exceeds unity. Experimental determination[8–10] of these yield functions, however, have given results too small by several orders of magnitude, particularly for the positive-ion yield per electron.

Very low gradient current-loading phenomena observed in Van de Graaff accelerator tubes have led to the postulation by Blewett[11] of a "Malter-layer" effect, and by McKibben and Boyer[10] of a positive-ion-nega-

* Work done under the auspices of the AEC.

[1] Fowler and Nordheim, Proc. Roy. Soc. (London) **A118**, 229 (1928).
[2] Stern, Gosling, and Fowler, Proc. Roy. Soc. (London) **A124**, 699 (1929).
[3] Hull and Burger, Phys. Rev. **31**, 1121(A) (1928).
[4] B. Snoddy, Phys. Rev. **37**, 1678(A) (1931).
[5] J. W. Beams, Phys. Rev. **44**, 803 (1933).
[6] A. J. Ahearn, Phys. Rev. **50**, 238 (1936).
[7] H. W. Anderson, Trans. Am. Inst. Elec. Engrs. **54**, 1315 (1935).
[8] J. G. Trump and R. J. Van de Graaff, J. Appl. Phys. **18**, 327 (1947).
[9] I. Filosofo and A. Rostagni, Phys. Rev. **75**, 1269 (1949).
[10] J. L. McKibben and K. Boyer, Phys. Rev. **82**, 315(A) (1951).
[11] J. P. Blewett, Phys. Rev. **81**, 305(A) (1951).

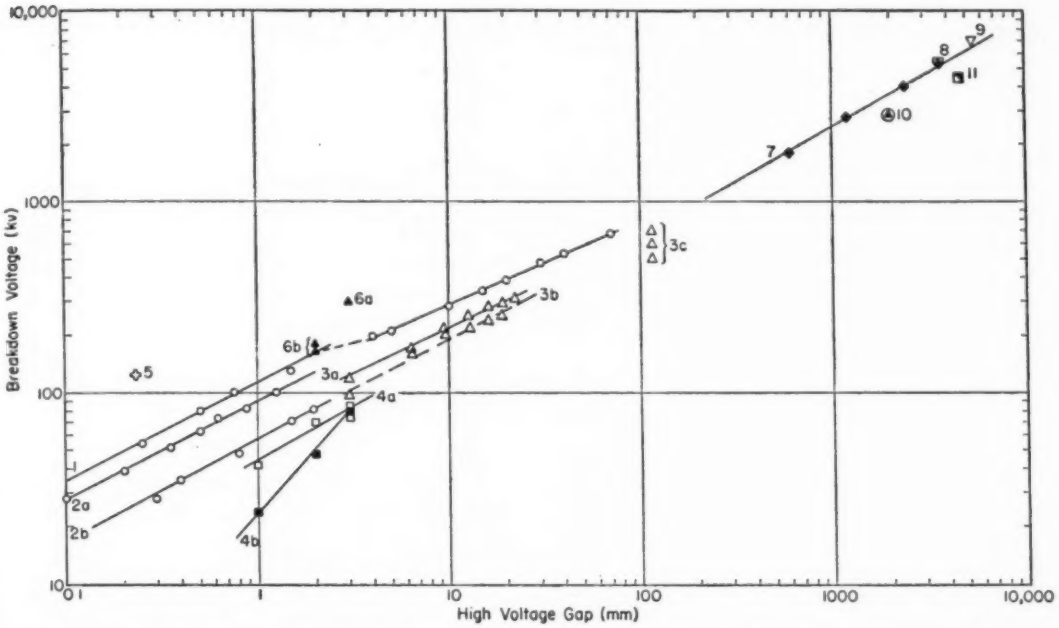


FIG. 1. Plot of data from the literature of breakdown voltage *vs* distance from highest to lowest potential electrode, for uniform-field and near-uniform-field geometry. Numbers on curves indicate sources as listed below.

(1) Trump and Van de Graaff, (see reference 8) 1-inch sphere steel anode, 2-inch steel disk, outgassed with glow discharge.
(2) Wm. Parkins, "Vacuum Sparking Potentials under Surge Conditions," MDDC 858, 18 February, 1946. Voltage applied in pulses of 3×10^{-7} sec rise time. (a) Tungsten hemispheres 2-inch diameter, out-gassed by spark discharge. (b) Copper hemispheres 2-inch diameter, out-gassed by spark discharge.
(3) J. L. McKibben and R. K. Beauchamp, "Insulation-Flashover Tests in Vacuum and Pressure," AECD 2039. (a) Flat aluminum. (b) Flat cold-rolled steel. (c) Van de Graaff test-section 4½ inches long, aluminum rings sandwiched between Mykroy rings, steel anode plate, negative end open to vacuum system in simulation of operation as beam tube, results on three test-sections.
(4) Gleichauf (see reference 12). (a) Kovar cathode, 18–8 st. steel anode, flat with rounded ends. (b) Copper, flat, with rounded ends (hole in center of anode).
(5) R. J. Piersol, British Assoc. Advancement of Science, Report 359 (1924). Molybdenum spheres after heating to 1400°C.
(6) J. L. Hayden, Am. Inst. Elec. Engrs. J. **41**, 852 (1922). (a) Molybdenum spheres 1-cm diameter out-gassed to red heat, polished. (b) Molybdenum spheres 1-cm diameter.
(7) Los Alamos big Van de Graaff, polished aluminum electrodes.
(8) Robinson *et al.*, Phys. Rev. (to be published).
(9) J. G. Trump (private communication on performance of new 12-Mev Van de Graaff machine, polished aluminum electrodes.
(10) Los Alamos small Van de Graaff machine (not limited by tube sparking), steel electrodes.
(11) Wisconsin Van de Graaff machine (not limited by tube sparking), steel electrodes.

tive-ion chain-reaction. This loading may be so severe that with power supplies of limited regulation, no breakdown can be observed. The relationship of these phenomena, and the mechanisms postulated to explain them, to breakdown itself it is not clear. In the case of the results obtained by McKibben and Boyer at least, it appears that loading and breakdown are due to different mechanisms, so that when the former is cleared up the latter has a chance to assert itself. This situation seem analogous to that with very short gaps, where ordinary field-emission is apparent, and the attainment of breakdown depends on sufficiently lowering the resistance of the power-supply, or clearing up the field emission by the smoothing of roughened surfaces. In what follows an attempt will be made to interpret the breakdown phenomenon for both long and short gaps in common terms, and in this picture the loading is incidental.

II. PROPOSED HYPOTHESIS AND SUPPORTING DATA

A hypothesis is proposed herein, the implications of which are briefly elaborated, and a summary of pertinent evidence is presented. The hypothesis is that the initiation of breakdown is due to detachment by electrostatic repulsion of a clump of material loosely adhering to one electrode, but in electrical contact with it; traversal by the clump of most or all of the high voltage gap, and impingement on an electrode at much lower, or at the lowest potential. It will be shown that this hypothesis and the experimental data are consistent with production by the clump of local temperatures in excess of any known boiling points. It will be assumed that production of such a condition will lead very quickly to the development of the full breakdown. Details of the subsequent events are excluded from the scope of this paper. Considering the scarcity of evidence regarding these details, there appears to be virtually no basis for preference among the various speculative possibilities. For the present it is not necessary even to specify on which electrode the clump is assumed to originate.

A quantitative formulation of this initiator hypothesis may be set down very simply as follows. Assume that breakdown will occur when the energy per unit area W delivered to the target electrode exceeds a value C', a constant, characteristic of a given pair of electrodes. This quantity W is just the product of the gap voltage

V by the charge density on the clump. The latter is proportional to the field E at the electrode of origin so that the breakdown criterion becomes simply

$$VE \geqq C, \tag{1}$$

where C is a product of C', some numerical factors, and possibly a field intensifying factory due to microscopic field inhomogeneities in the neighborhood of the clump during detachment from its parent electrode. For the case of plane-parallel electrodes where $E=V/d$, this criterion predicts at once that the voltage which a gap can sustain is proportional to the square-root of the gap length for a given pair of electrodes.

$$V \geqq (Cd)^{\frac{1}{2}}. \tag{2}$$

The prediction of this simple result has led to search of the literature for pertinent evidence, a summary of which is given in Fig. 1 in the form of a double-log plot of voltage *vs* length of gap, for uniform and nearly uniform field conditions. The only contradictory evidence is that provided by a three-point curve given by Gleichauf[12] for copper electrodes in the range 1 to 3 mm, whereas all other data are consistent with this prediction. Experimental data reported in the literature are usually given in the form of voltage-gap plots on linear scales, without further analysis.[13] The most striking features of Fig. 1 are the clustering of the data about straight lines of slope $\frac{1}{2}$, and the range of 10^5 in gap over which the relationship is approximately valid. Also noteworthy is the fact that the data of the various investigators, most of whom have taken pains to obtain conditions favorable to high voltage, do not vary greatly from one another and that a typical value for the constant C is about 3 megavolts squared per foot.

It is clear that if the voltage-field product plays the significant role assigned to it here, the single quantity C suffices as a figure of merit for the vacuum insulating qualities of a given gap, and provides a basis for comparison of data taken under different conditions of gap and voltage.

The magnitude of C is clearly a function of the condition of the electrodes. It is interesting to note that points 5 and 6*a* in Fig. 1 were obtained by different workers for well-outgassed molybdenum, and correspond to values of C of 20 and 6 megavolts squared per-foot, whereas points 6*b* represent molybdenum not outgassed.

Using the foregoing data, one may attempt a crude estimate of temperatures produced by a clump traversing a gap, as follows. The energy W' delivered by the clump per atom struck on the surface of the target electrode is just

$$W' \geqq Ca^2/4\pi, \tag{3}$$

where a is the interatomic distance. If n is the depth of electrode, expressed in number of atom layers which share the energy of the clump, the local temperature attained is

$$T \geqq Ca^2/4\pi nk, \tag{4}$$

where k is Boltzmann's constant. From the empirical data above C is typically 10^6(*ab*volts^2/cm), equivalent to 2.7 megavolts^2/ft, which gives, for $n=1$, $T \geqq 10^6$ °K, or disruptive temperatures (i.e., temperatures above metallic boiling points) may be expected for penetrations of several hundred atomic layers.

The impact velocity v on the assumption of spherical clumps of radius r whose density is of order unity is given by

$$v \geqq [C/8r]^{\frac{1}{2}}. \tag{5}$$

For $r \geqq 10^{-2}$ to 10^{-6} cm, $v \geqq 10^4$ to 10^6 cm/sec. These estimates imply that breakdown due to small clumps could be effective for the full gap-voltage up to megacycle frequencies for short gaps. These data therefore suggest that one way of obtaining further evidence regarding high voltage breakdown in vacuum is to study the effect of frequency of applied voltage on the relation between breakdown-voltage and gap-distance.

One pertinent bit of evidence in this connection is the result reported by Halpern *et al.*[14] that at 2800 mc/sec, 2.0 megavolts were supported across a 2.0-inch gap, or 24 megavolts^2/per foot, presumably across unoutgassed copper. This is much higher than has ever been obtained at dc with copper and is consistent with the fact that clump breakdown for the full gap-voltage is here precluded by transit time considerations. At present there appears to be no firm evidence of a "total-voltage" effect under conditions wherein clump breakdown at full voltage is precluded by transit-time considerations.

The data of Fig. 1 are pertinent to the special case of uniform geometry. The question naturally arises as to how well Eq. (1) is satisfied for more general cases. Relevant to this question are the data of the upper part of curve 1 in Fig. 1. The data for this curve were taken with a 1-inch diameter spherical anode and plane cathode for the low voltage region, the extension to higher voltage requiring the insertion along the gap of plane electrodes to control the potential distribution. Thus the gradient is not accurately calculable for the higher voltage range, though it is reasonable to assume the deviation from uniformity is not great. It is interesting to note, however, that the high voltage extension of curve 1 has a slope definitely smaller than 0.5 (almost 0.43), and this deviation is in the proper direction if either electrode gradient enters into the picture, as assumed in Eq. (1).

Other data relevant to this question are summarized in Table I. It is clear that these are too limited to be conclusive. The spread of the results is consistent

[12] P. Gleichauf, J. Appl. Phys. **22**, 766 (1951).

[13] It is understood via private communication from E. Bretscher and J. L. McKibben that a summary of data similar to Fig. 1 has been assembled by R. L. Fortescue of the British Atomic Energy Research Establishment at Harwell.

[14] Halpern, Everhart, Rapuano, and Slater, Phys. Rev. **69**, 688A (1946).

TABLE I.

Reference	Geometry	Anode	Cathode	Electrode treatment	V(10^6 volts)	$E_{cathode}$ 10^6 volts/ft	$VE_{cathode}$ 10^{12} volts^2/ft
C. C. Chambers J. Franklin Inst. 218, 463 (1934)	Cylindrical	Nickel	0.5 Mil tungsten	Wire heated to 2800°K for 30 min.	0.011±10%	78	0.86
			1.6 Mil tungsten	Cylinder heated to dull red heat for 5 min.	0.023±10%	60	1.30
A. J. Ahearn, reference 6 and Phys. Rev. 44, 277 (1933)	Cylindrical	2 cm diam nickel	Thoriated W 1.0 Mil		0.017	60	1.0
			Thoriated W 1.0 Mil		0.024	78	1.9
			Pure W 0.6 Mil		0.015	84	1.3
			Pure W 1 Mil		0.023	81	1.9
R. C. Mason, Phys. Rev. 52, 126 (1937)	Line-plane	2″-diam. Cu disk	28 mil tungsten wire in 1″-diameter seimi-circle	Everything baked out to 500°C	0.035	18	0.63
K. Hashimoto, reference 15	Point-plane	Nickel Plane	Nickel point	Conditioned by sparking	0.030 } for 1 mm separation		
		Nickel Point	Nickel plane		0.090 }		

equally with an interpretation in terms of cathode gradient, and in terms of voltage-field product. Applicability of the proposed theory to these cases is suggested, however, by the fact that values of C are in the same range as the data of Fig. 1. If the mechanism for these cases is indeed the same as for the uniform-field, high voltage gap, one is almost obliged to say for these cases, more specifically than has been said thus far, that it is the product of the total voltage by the *cathode* gradient which controls the beginning of the discharge. Hashimoto's data,[15] though they do not allow the calculation of gradients, indicate that breakdown for a high gradient anode is of the same order of magnitude as for a high gradient cathode, other things being the same, and may therefore be interpreted in terms of a similar mechanism. The fact that it is higher is in agreement with many other reported qualitative observations. This suggests that, generally, the gradient which controls breakdown, and the place of origin of the hypothetical clump, is at the cathode, subject to the stipulation that the anode is not at substantially higher gradient, in which case it dominates.

In addition to the voltage-field product, certain qualitative aspects of vacuum breakdown can be plausibly explained by the initiation mechanism postulated here. The so-called conditioning process is well known, whereby maintenance of a gap at high voltage, with or without occasional sparking, improves the insulation strength of the gap. One surmises, on the clump theory, that conditioning is simply the process of detachment of the most loosely adhering material, and its firm embedment by acceleration across the gap. Direct evidence for such material transfer is cited by Anderson who used copper and steel for anode and cathode, respectively, and reported deposition of a brownish deposit on the cathode before any breakdown had occurred, with 20 minutes steady application of high voltage. After breakdown sufficient material was transported to provide direct spectroscopic evidence of transfer from the anode. No data are given on transfer from the cathode.

Another well-known fact about vacuum breakdown is its sporadic character. Thus, a gap may hold voltage for minutes or hours, then suddenly break down, and there is no evidence reported of current build-up, or other continuous cumulative effect prior to breakdown. Such a characteristic seems quite in accordance with a hypothesis which postulates a single event as initiator. The difficulty of reproducing breakdown voltages, even for a given gap, is also in agreement with the general character of a clump mechanism. It may be added that at least some of the electronic loading referred to in the introduction, particularly that which appears close to the breakdown voltage, may be due to clump-initiated cathodic emission corresponding to a voltage-field product less than the critical value C.

From the practical point of view of improving the high voltage insulating characteristics of vacuum, experience agrees with what one might expect from the predictions of the hypothesis. Elimination of loose material, the first obvious step, need hardly be recommended to the experimenter in this field. Presumably, measures which have proven successful in the past, such as cleaning, choosing proper materials, and high temperature bake-out either have progressively removed or consolidated the loose material, or have increased the threshold temperature for the development of breakdown.

Very recently quantitative data of a sort similar to that given in Fig. 1 have become available[13] on breakdown in a vacuum gap of variable length bridged by an insulator. These are replotted in Fig. 2, for the one case known, a Pyrex cylinder between flat copper electrodes. The slope of the straight line obtained again

[15] K. Hashimoto, J. Phys. Soc. Japan 2, 71 (1947).

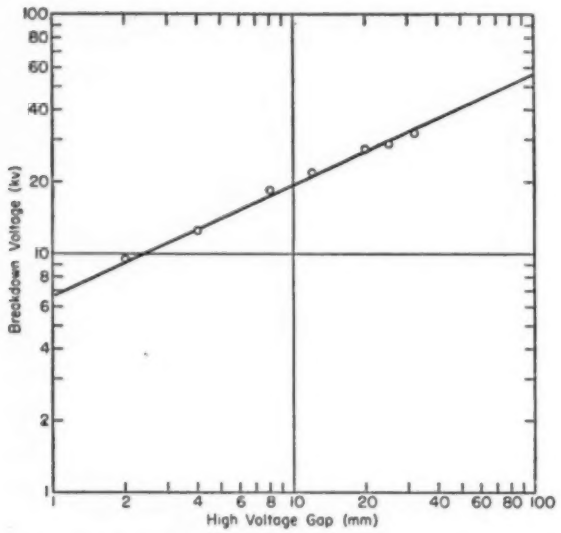


FIG. 2. Replot of data from Gleichauf (see reference 12) of breakdown-voltage *vs* gap-distance for Pyrex cylinder between flat copper electrodes.

is 0.5, but the value of C is here only 0.1 megavolts squared per foot. This is in accord with the well-known result that, at least in short lengths, insulators in vacuum are inferior to the vacuum gap between metal electrodes themselves. Thus, all the data on very short gaps in Fig. 1 were obtained with an arrangement that provided a much longer breakdown path along the insulator than across the gap itself. For the larger gaps the length of insulator required becomes troublesomely long. The now well-known device for circumventing this difficulty is to segment the insulator, arranging in effect a sandwich of alternating insulators and conductors, the latter tied to a potential dividing system. The data of Fig. 1 given by McKibben, and by Trump and Van de Graaff were obtained with two- and four-decker sandwiches, respectively. Modern high voltage x-ray and Van de Graaff accelerator tubes represent a large scale extension of this principle. It is interesting to note that the data of Figs. 1 and 2 together provide one with a guide for estimating the performance of Van de Graaff accelerator tubes as a function of their length, and provide the sort of data necessary to determine in a systematic way the segmentation interval for the tube and potential dividing system.

III. CONCLUSION

The results given here should be useful for practical purposes independently of the validity of the clump hypothesis itself. In addition, these results provide a quantitative criterion for the appraisal of any proposed theory of the initiation mechanism. In fact the strongest single reason for considering the clump theory is that, at present, it is the only one available which predicts the voltage-field constancy without artificial, *ad hoc* assumptions.

The author is pleased to acknowledge the encouragement and stimulation provided by discussions with Professor Ray G. Herb of the University of Wisconsin and Dr. Keith Boyer of this Laboratory.

Conference on Medium Energy Nuclear Physics

The University of Pittsburgh is planning to have a Conference on Medium Energy Nuclear Physics (in Pittsburgh, Pennsylvania) on June 5, 6, and 7. The program will be limited to studies of nuclei with $A>4$, using projectiles accelerated in electrostatic generators or conventional cyclotrons.

Eddy-Currents in Solid Cylindrical Cores Having Non-Uniform Permeability

HAROLD ASPDEN
Trinity College, Cambridge, England
(Received December 12, 1951)

A method of estimating magnetization losses due to eddy-currents in solid cylindrical cores of ferromagnetic material is given. The conception of complex permeability is introduced to show the effects which hysteresis has upon the degree of flux penetration. A formula is given for the effective permeability which should be used in the classical treatment in which nonlinearities have been ignored.

An arbitrary case in which the permeability changes across the core section is considered and the solution is extended to a core which is homogeneous, but whose permeability changes with the degree of magnetization. In this way errors due to a nonlinear $B-H$ curve are reduced and the complex permeability notation is shown to lead to a more reasonable Steinmetz relationship.

The eddy-current anomaly is mentioned and it is noted that the theory developed is suited to a consideration of the anomaly as explained in terms of time-lag effects.

INTRODUCTION

WHEN calculating eddy-currents in a ferromagnetic system it is usual to assume that the permeability does not change with time and is the same at all points in the system. Although in many instances the latter assumption may well be justified, the former assumption, which ignores hysteresis, is seldom justified. A mathematical approach to the problem of varying permeability is afforded by the conception of a complex value of permeability. This is equivalent to presuming that the hysteresis loop is elliptical in shape, or that there is a definite time lag between the magnetizing field and the flux density at a particular point. In so far as hysteresis loss is concerned, such an assumption is only by way of an approximation to the shape of the hysteresis loop since hysteresis loss is not the result of time-lag effects. However, that such time-lag effects do exist is becoming increasingly evident since they afford explanations for the residual magnetization losses in ferrites and the phenomenon of the "extra loss" in metallic ferromagnetic materials which is generally termed the eddy-current anomaly. It is the present purpose to calculate the loss associated with the eddy currents induced in a solid cylindrical core having a complex permeability and to interpret the physical significance of the result.

A problem of practical importance is also presented by the case where the permeability varies across the section of the core. Under conditions where there are appreciable magnetic skin effects most of the magnetic flux is concentrated near the surface and most of the eddy loss occurs there. If the permeability at the surface is substantially different from that well within the core the normal calculations which are based upon a uniform permeability will not apply, but the effects which such permeability variation has upon the eddy loss is of interest. Little useful purpose will be served by attempting an analysis with a general form of permeability distribution. The problem has, therefore, been restricted to the consideration of a particular case in which the permeability varies in a certain mathematically convenient manner.

THE GENERAL EQUATION

The core is assumed to be very long, solid, and of cylindrical shape, the radius being a. Let the permeability be the same at all points equidistant from the central axis of the core and have a value μ at a radius x. Then, if H is the magnetic field directed along the core at radius x, the total flux linking an elemental ring of radius r is $2\pi\int_0^r \mu x H dx$. Hence the electromotive force e induced in the ring is (see Fig. 1)

$$e=-2\pi\frac{\partial}{\partial t}\int_0^r \mu x H dx. \tag{1}$$

The resistance of the ring to circulating currents is $2\pi r/\sigma dr$ per unit length, dr being the thickness of the ring, and σ the specific conductivity of the material of the core. If i is the current density at r, the current in the ring per unit length of core is

$$i dr=(e/2\pi r)\sigma dr. \tag{2}$$

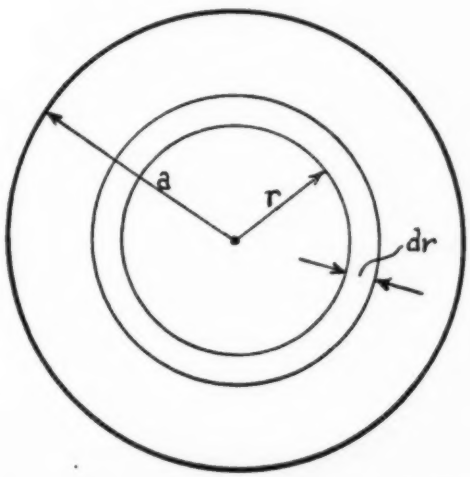


FIG. 1. Core section showing elemental ring of eddy-current path.

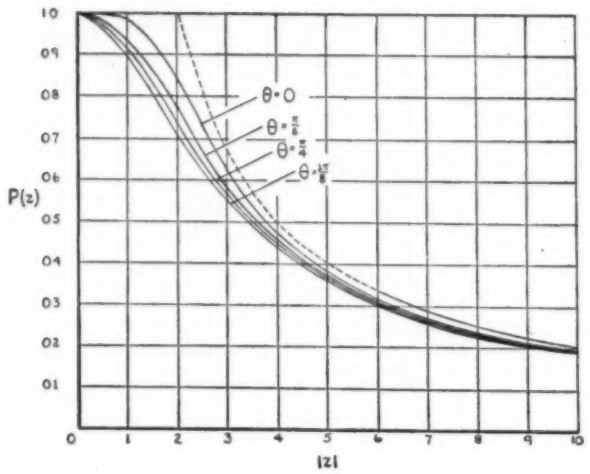


FIG. 2. The penetration function.

From (1) and (2),

$$i=-(\sigma/r)(\partial/\partial t)\int_0^r \mu x H dx. \tag{3}$$

The magnetizing field within the ring caused by the eddy current circulating in it is $-dH=4\pi i dr$ so that

$$i=-(1/4\pi)(dH/dr). \tag{4}$$

After equating (3) and (4),

$$r(dH/dr)=4\pi\sigma(\partial/\partial t)\int_0^r \mu x H dx. \tag{5}$$

Differentiation of this equation with respect to r and insertion of the limits gives the general equation

$$rd^2H/dr^2+dH/dr-4\pi\sigma r\partial/\partial t(\mu H)=0. \tag{6}$$

When H is presumed to be varying in a purely sinusoidal manner with respect to time, the angular velocity being ω, $\partial H/\partial t$ may be expressed as $j\omega H$, so in the instances where μ may be taken to be a constant with respect to time Eq. (6) reduces to

$$rd^2H/dr^2+dH/dr-j4\pi\sigma\omega\mu rH=0. \tag{7}$$

SOLUTIONS FOR COMPLEX PERMEABILITY

It is presumed that $\mu=\mu_0\epsilon^{-j\theta}$, where μ_0 and θ are constants. The physical significance of this will be discussed later, and it will be shown that while the permeability as defined by the ratio of the flux density to the field strength is by no means constant when μ is as above, it may be regarded as a constant vector quantity for the purposes of solving Eq. (7). In this solution attention has been restricted to the case in which purely sinusoidal variations of H are concerned, the angular velocity being ω.

Substitution of this complex value of μ in Eq. (7) gives

$$rd^2H/dr^2+dH/dr-j4\pi\sigma\omega\mu_0\epsilon^{-j\theta}rH=0. \tag{8}$$

After putting $z^2=4\pi\sigma\omega\mu_0a^2\exp\{-j[(\pi/2)+\theta]\}$, this becomes

$$rd^2H/dr^2+dH/dr+r(z/a)^2H=0. \tag{9}$$

The general solution of this equation is

$$H=AJ_0(zr/a)+BY_0(zr/a), \tag{10}$$

where A and B are constants, and J_0 and Y_0 are Bessel functions of zero order and of the first and second kinds, respectively. Since $Y_0(0)$ is $-\infty$, and we know H to be finite on the axis of the core, $B=0$. Hence, if $H=H_0$ at the surface of the core where $r=a$,

$$H=H_0J_0(zr/a)/J_0(z). \tag{11}$$

ϕ, the total flux within the core, is obtained by integrating μH over the cross section of the core, thus

$$\phi=2\pi\mu_0\epsilon^{-j\theta}\int_0^a rHdr. \tag{12}$$

Now, noting that $\int rJ_0(kr)dr=1/krJ_1(kr)$, where k is a constant and J_1 is a Bessel function of the first kind and order one, we have, from Eqs. (11) and (12)

$$\phi=2\pi\mu_0\epsilon^{-j\theta}H_0a^2/zJ_1(z)/J_0(z). \tag{13}$$

We are concerned with the degree to which the flux penetrates into the core and the magnetization loss. The penetration function is defined as the ratio which the maximum value of the mean flux density within the core bears to the maximum flux density at the surface. Call the function P, and we have

$$P=1/\pi a^2\frac{|\phi|}{|\mu_0\epsilon^{-j\theta}H_0|}, \tag{14}$$

or

$$P(z)=2/|z|\frac{|J_1(z)|}{|J_0(z)|}. \tag{15}$$

z is a complex quantity which may be expressed in the form $z=|z|\exp\{j[(3\pi/4)-(\theta/2)]\}$, and the function P is shown in Fig. 2 for various values of θ. The dotted line shows the hyperbola $P=2/|z|$ to which the curves tend when $|z|$ is very large. The magnetization loss can be calculated directly from Eq. (13). The rate of loss per unit length of core is the mean value of the real component of $(H_0/4\pi)(\partial\phi/\partial t)$. Call this loss W, and we have

$$W=R_e\left[\frac{j\omega\phi}{4\pi H_0}H_m^2\right], \tag{16}$$

where H_m is the root-mean-square value of H_0. From Eqs. (13) and (16)

$$W=R_e\left[-\frac{\omega\mu_0a^2}{2}\exp\{-j[(\pi/2)+\theta]\}\frac{H_m^2J_1(z)}{zJ_0(z)}\right], \tag{17}$$

or, noting that $\omega\mu_0 a^2 \exp\{-j[(\pi/2)+\theta]\}=z^2/4\pi\sigma$,

$$W=H_m^2/8\pi\sigma R_e\left[-z\frac{J_1(z)}{J_0(z)}\right]. \quad (18)$$

It is more usual to consider the loss per unit volume $W/\pi a^2$ or W_v. From (17) we have

$$W_v=\omega\mu_0/2\pi H_m^2 L(z), \quad (19)$$

where

$$L(z)=R_e\left[\exp\{j[(\pi/2)-\theta]\}\frac{J_1(z)}{zJ_0(z)}\right]. \quad (20)$$

The loss function L is shown graphically in Fig. 3. The dotted lines depict the asymptotes $1/|z|\ \sin(\pi/4+\theta/2)$ to which the curves tend as $|z|$ increases. In the case where $|z|$ is small L may be calculated from the following series:

$$L(z)=1/2\sin\theta+1/16|z|^2\cos2\theta-1/96|z|^4\sin3\theta -11/6144|z|^6\cos4\theta. \quad (21)$$

SOLUTIONS FOR VARIABLE PERMEABILITY

When the permeability varies across the section of the core the problem is much more complicated. Equation (7) can be solved for any particular expression of the form $\mu=f(r)$ by expanding the differential equation as a power series and equating the coefficients of powers of r. Alternatively, it may happen that some simple solution is possible. This is very unlikely, but there is a group of variable permeability distributions which are susceptible to ready solution. These will now be considered and the significance which may be attached to them will be given later.

It is assumed that the permeability may be expressed as $\mu=\mu_0 e^{-j\theta}(r/a)^p$, the complex conception of permeability still being retained. In this case it is convenient to put $z^2=(2/2+p)^2 4\pi\sigma\omega\mu_0 a^{2-p}\exp\{-j[(\pi/2)+\theta]\}$, and Eq. (7) then becomes

$$r(d^2H/dr^2)+dH/dr+(2+p/2)^2(z/a)^2 r^{p+1}H=0. \quad (22)$$

It is now a simple matter to show that if $y=r^{1+p/2}$ (22) becomes

$$yd^2H/dy^2+dH/dy+y(z/a)^2H=0, \quad (23)$$

which is similar to (9) and so has a similar solution. However, at the surface of the core $y=a^{1+p/2}$ so that the solution is

$$H=H_0J_0(zy/a)/J_0(za^{p/2}). \quad (24)$$

ϕ, the total flux within the core, is obtained by integrating μH over the cross section of the core, thus:

$$\phi=2\pi\mu_0 e^{-j\theta}\int_0^a r(r/a)^p H dr, \quad (25)$$

or

$$\phi=2\pi(2/2+p)\mu_0/a^p e^{-j\theta}\int_0^{a1+p/2} yH dy. \quad (26)$$

Combine (24) and (26), and we have

$$\phi=2\pi\left(\frac{2}{2+p}\right)\frac{\mu_0}{a^p}\frac{H_0 e^{-j\theta}}{J_0(za^{p/2})}\int_0^{a1+p/2} yJ_0(zy/a)dy. \quad (27)$$

Evaluating this integral leads to the expression

$$\phi=2\pi\left(\frac{2}{2+p}\right)\frac{\mu_0}{a^{p/2-2}}\frac{H_0 e^{-j\theta}}{z}\frac{J_1(za^{p/2})}{J_0(za^{p/2})}. \quad (28)$$

A penetration function will be of little interest in this case so attention will be restricted to the determination of the magnetization loss. From (16) and (28) the loss per unit length of core is given by

$$W=R_e\left[\frac{\omega\mu_0}{2+p}a^{2-p/2} \times\exp\{j[(\pi/2)-\theta]\}\frac{H_m^2 J_1(za^{p/2})}{zJ_0(za^{p/2})}\right]. \quad (29)$$

Hence the loss per unit volume is

$$W_v=\frac{W}{\pi a^2}=\frac{\omega\mu_0 H_m^2}{2\pi}\left(\frac{2}{2+p}\right) \times R_e\left[\exp\{j[(\pi/2)-\theta]\}\frac{J_1(za^{p/2})}{za^{p/2}J_0(za^{p/2})}\right]. \quad (30)$$

This reduces to

$$W_v=\omega\mu_0 H_m^2/2\pi(2/2+p)L(za^{p/2}), \quad (31)$$

where L is the same loss function as is shown in Fig. 3, it being noted that z is defined in a slightly different way. A further simplification is possible by removing the factor p and replacing it in the appropriate way by the mean permeability $\mu_m e^{-j\theta}$. The mean permeability is

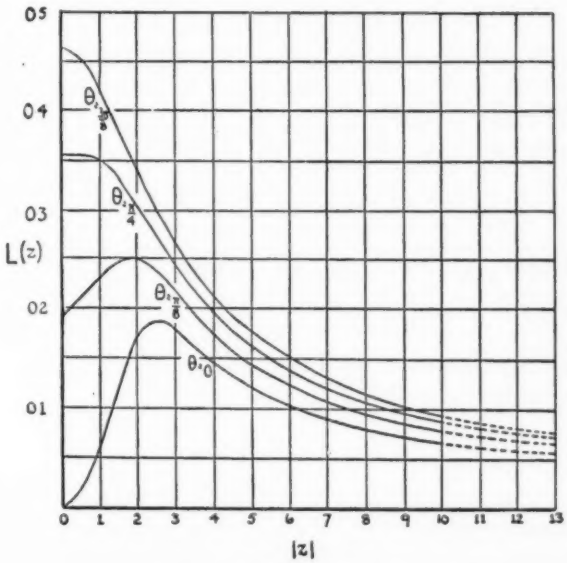


FIG. 3. The loss function.

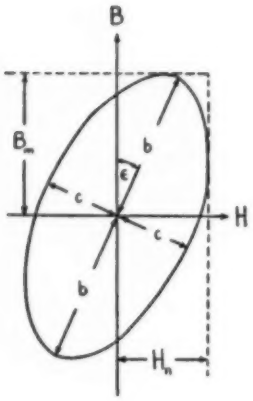


FIG. 4. Elliptical $B-H$ relationship depicted by complex permeability notation.

given by

$$\mu_m e^{-j\theta} = 2/a^2 \int_0^a \mu r dr$$

$$= 2\mu_0 e^{-j\theta}/a^{2+p} \int_0^a r^{1+p} dr$$

$$= \mu_0 e^{-j\theta}(2/2+p). \tag{32}$$

From (31) and (32) we have

$$W_v = \omega\mu_m/2\pi H_m^2 L(za^{p/2}), \tag{33}$$

which compares with (19). In this case the factor $za^{p/2}$ is given by

$$za^{p/2} = (2/2+p)a(4\pi\sigma\omega\mu_0)^{\frac{1}{2}} \exp\{j[(3\pi/4)-(\theta/2)]\}, \tag{34}$$

which, from (32), becomes

$$za^{p/2} = \mu_m a(4\pi\sigma\omega/\mu_0)^{\frac{1}{2}} \exp j(3\pi/4-\theta/2). \tag{35}$$

This compares with the value of z before which could be expressed as

$$z = \mu_0 a(4\pi\sigma\omega/\mu_0)^{\frac{1}{2}} \exp\{j(3\pi/4)-(\theta/2)\}. \tag{36}$$

The above constitutes a very simple solution of an arbitrary system involving non-uniform permeability and it now remains to consider the significance of the results

COMPLEX PERMEABILITY

In the case of purely sinusoidal variations of H with respect to time the conception of a constant complex value of permeability implies a fixed angular displacement between the related flux density and magnetizing field. By taking $\mu = \mu_0 e^{-j\theta}$, this may be expressed mathematically by the two equations

$$H = H_n \sin\omega t, \tag{38}$$

$$B = B_m \sin(\omega t - \theta). \tag{39}$$

B is the flux density. By eliminating the parameter ωt from (38) and (39) the relation between B and H is an ellipse as shown in Fig. 4. It may be shown that the values of b, c, and ϵ as depicted in the Fig. 4 are given by

$$b \simeq \mu_0 H_n, \tag{40}$$

$$c \simeq H_n \sin\theta, \tag{41}$$

$$\epsilon \simeq \cos\theta/\mu_0. \tag{42}$$

It is evident that the shape of the loop is independent of ω, i.e., the frequency and the area are

$$\pi bc \simeq \pi\mu_0 H_n^2 \sin\theta. \tag{43}$$

Since B_m, the maximum flux density, is approximately $\mu_0 H_n$, the area is proportional to $B_m^2 \sin\theta/\mu_0$. By taking μ_0 and θ as constant we are simulating the case in which the area of the loop is proportional to the square of the maximum flux density. To simulate hysteresis loss the area of the loop has to be independent of frequency and has to be proportional to a power of the maximum flux density of the order of 1.6. The use of complex permeability affords a valuable means of allowing for the hysteresis loss since a consideration of the actual permeability as deduced by dividing (39) by (38) shows that it may vary from zero to infinity, whereas when considered as a vector quantity the permeability deduced in this way is a constant independent of time. For the purposes of Eq. (7) there is no reason why μ cannot be expressed as a vector quantity.

An actual hysteresis loop has the form shown in Fig. 5. The irregular shape leads to the production of harmonics in the flux density and magnetizing field wave forms, but these are ignored in the present treatment. The most appropriate choice of μ_0 and θ for a given hysteresis loop comprises the determination of the ellipse (shown dotted in Fig. 5) which represents the same loss and loss angle. θ is the loss angle of the ellipse and so is chosen to be equal to that of the hysteresis loop. If A is the area of the hysteresis loop, then μ_0 is chosen so that its value is

$$\mu_0 = \pi B_m^2 \sin\theta/A. \tag{44}$$

The combined effects of eddy loss and hysteresis loss may be considered by this method of analysis. In Fig. 2, the curve with $\theta=0$ corresponds to the case where there is no hysteresis loss and the permeability is constant, whereas the other curves apply to cases where hysteresis loss is present. It will be noted that the degree of penetration is very much the same whether hysteresis loss occurs or not, providing, of course, $|z|$ is the same. For the mathematical treatment which neglects hysteresis to be reasonably valid, in spite of the presence of hysteresis, the effective permeability must be chosen in accordance with Eq. (44) above. The magnetization loss function for large $|z|$ tends to $1/|z| \sin[(\pi/4)+(\theta/2)]$. This would imply that at high frequencies the loss is very much dependent upon θ. Since θ is a factor associated with the hysteresis loss it may appear that hysteresis loss will form a major part of the total magnetization loss. Actually, the loss is mainly eddy loss. This apparent anomaly arises from the factor $|z|$

For the purposes of loss calculation we may take the permeability as $\mu_0 e^{-i\theta}/(1+\sin\theta)$. If z_e is the value of z calculated by using this value of permeability we have

$$z_e = z/(1+\sin\theta)^{\frac{1}{4}}. \tag{45}$$

Consequently, the magnetization loss function tends to the value

$$L(z) = \frac{1}{|z_e|}\frac{\sin(\pi/4+\theta/2)}{(1+\sin\theta)^{\frac{1}{4}}} = \frac{1}{|z_e|}\sin\frac{\pi}{4}, \tag{46}$$

which is the value calculated when hysteresis loss is ignored. An important result of this is that at high frequencies the calculations of magnetization loss based upon negligible hysteresis and uniform permeability may be considered to be substantially representative in spite of these foundations providing the effective permeability is taken as

$$\mu_0 = \pi B_m^2 \sin\theta/A(1+\sin\theta). \tag{47}$$

At low frequencies it is of interest to consider the combined effects of eddy and hysteresis losses in view of the eddy current anomaly. It is a well-known experimental fact that in many ferromagnetic materials there is a greater frequency-dependent component of the magnetization loss per cycle than can be accounted for by normal eddy-current calculations. Equation (21) gives $L(z)$, a function which is proportional to the magnetization loss per cycle. $|z|^2$ is proportional to frequency, and evidently the term $1/2\sin\theta$ would correspond to hysteresis loss and $1/16|z|^2\cos\theta$ would represent the eddy loss component at very low frequencies. It might appear that the eddy loss is dependent upon θ but this may be accounted for by reference to Eq. (19). Low frequency eddy loss considerations are usually based upon the assumption that the range of change of the total flux within the core is independent of frequency. For this to be so H_m must change with frequency, and it may be shown that H_m is proportional to $1/P(z)$. It may also be shown that

$$P(z) = 1 - |z|^2/8\sin\theta\cdots. \tag{48}$$

Accounting for this change in H_m we note that the loss for constant total flux density conditions is proportional to $L(z)H_m^2$ or $L(z)/P^2(z)$, whence we have

$$L(z)H_m^2 \propto \frac{1/2\sin\theta+1/16|z|^2\cos2\theta\cdots}{(1-1/8|z|^2\sin\theta\cdots)^2} \propto 1/2\sin\theta+1/16|z|^2\cdots. \tag{49}$$

Consequently, the eddy loss under these conditions may be regarded as being independent of θ. This shows that the eddy-current anomaly cannot be regarded as being a effect directly due to hysteresis. Brailsford[1] has shown that the eddy loss may be larger due to the wave-form distortion produced by the nonlinear shape of the hysteresis loop but this effect is insufficient to account for the increase in loss observed. Recently, Stewart[2] has shown that there is a time-lag in the build-up of the magnetic flux following the application of the field. Such a time-lag can account for the anomaly loss. The theory developed is ideally suited to the case where there is such a time constant since it leads to an elliptical magnetization loop. If τ is this constant, we have, putting $\theta=\omega\tau$ in Eq. (49)

$$L(z)H_m^2 \propto 1/2\sin\omega\tau+1/16|z|^2+\cdots. \tag{50}$$

If $\omega\tau$ is small this first term will itself give a frequency-dependent component which will be similar to the eddy term $1/16|z|^2$ since they are both directly proportional to ω.

VARIABLE PERMEABILITY

Even though the core be homogeneous in character it does not follow that we can assume the permeability to be uniform. When the core is subjected to cyclical magnetization, the shielding effects of the induced eddy-currents result in the range of change of the magnetizing field decreasing from the surface of the core to the center. In practice it is known that a magnetic material has a smaller effective permeability at a low flux density than at a high flux density. For a given material, there is a definite curve relating the effective value of the permeability and the range of change of magnetizing field. To attempt the solution of a problem when such complex considerations are involved is not worth while unless many simplifications and approximations are made. Let us consider the problem in which we are given a definite curve as above and require to determine the magnetization loss within a particular core at a given frequency and for a given applied magnetizing field. These conditions will determine the manner in which H, the magnetizing field, decreases from the core surface to the center. The permeability *versus* H curve combined with this will determine a definite decrease in permeability from the surface of the core to the center.

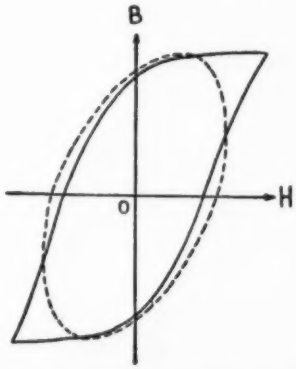


FIG. 5. Hysteresis loop and equivalent ellipse.

[1] F. Brailsford, J. Inst. Elec. Engrs. 95, Pt. II, 38 (1948).
[2] K. H. Stewart, Proc. Inst. Elec. Engrs. 97, Pt. II, 121 (1950).

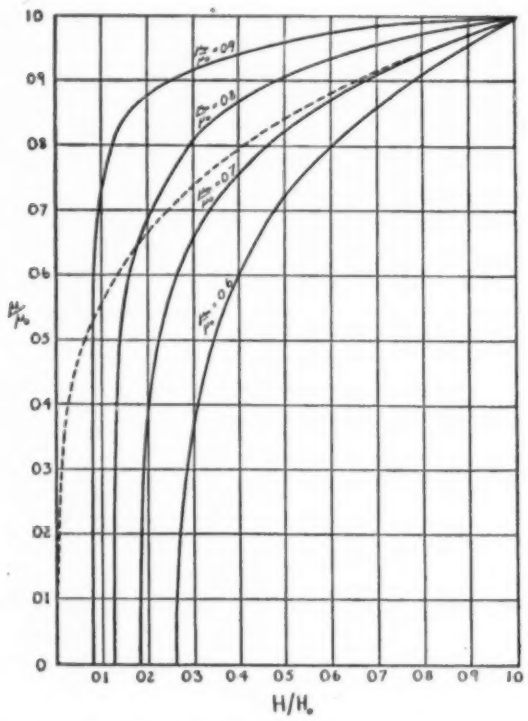


FIG. 6. $\mu-H$ relationships of an homogeneous core which are encountered with the type of sectional variation of effective permeability chosen.

The principal assumption to be made is that this decrease can be represented by a value of permeability given by $\mu e^{-i\theta}=\mu_0(r/a)^p e^{-i\theta}$. Each different value of the factor p corresponds to a different permeability *versus* H curve and the problem is to determine the particular value of this factor which best suits the actual form of this curve. Once this factor is known, the magnetization loss can be calculated from the formulas already derived.

Adopting the same terminology as before, and Eq. (24) gives the relationship between H and r

$$H=H_0\frac{J_0(zr^{1+p/2}/a)}{J_0(za^{p/2})}. \tag{51}$$

The values of a, $\sigma\omega$ and the range of H_0 having been selected, the appropriate value of μ_0 and θ as they apply at the surface for this particular H_0 range are determined. It is assumed that θ is constant. Equation (35) enables us to calculate k, where

$$za^{p/2}=\mu_m/\mu_0 k. \tag{52}$$

Also from Eq. (32),

$$1+p/2=\mu_0/\mu_m. \tag{53}$$

After combining (51), (52), and (53) we get

$$H/H_0=\frac{J_0\left(\frac{\mu_m}{\mu_0}k[r/a]^{\mu_0/\mu_m}\right)}{J_0\left(\frac{\mu_m}{\mu_0}k\right)}. \tag{54}$$

Finally, we may replace r in this equation by introducing μ, the effective permeability. This gives a relationship between μ and H in terms of the parameter μ_m, thus

$$H/H_0=\frac{J_0\left(\frac{\mu_m}{\mu_0}k\left[\frac{\mu}{\mu_0}\right]^{\mu_0/2(\mu_0-\mu_m)}\right)}{J_0\left(\frac{\mu_m}{\mu_0}k\right)}. \tag{55}$$

Figure 6 shows this relationship for the particular case when $k=5$ and $\theta=\pi/4$. These curves must then be compared with the known μ, H relationship and the most appropriate value of μ_m can be deduced by direct comparison. For the purposes of loss calculation it is desirable that the chosen curve compares well with the actual curve over the range representing the larger values of H.

By allowing for the decrease in permeability in this way the hysteresis effects are simulated in a more representative manner. From Eq. (43) the hysteresis loop has an area $\pi\mu H_n^2\sin\theta$, which may be written as $\pi B_m^2\sin\theta/\mu$. Since μ varies as the range of H_n or B_m in accordance with the above curves it is evident that the area of the loop can be represented as proportional to a power of B_m. The dotted curve in Fig. 6 has the equation

$$\mu/\mu_0=(H/H_0)^{0.25}, \tag{56}$$

and it will be noted that a curve of this form can closely represent the other calculated curves of Fig. 6 over the major and most important part of the range of H. In this particular instance the area of the hysteresis loop would be proportional to $B_m^2/B_m^{0.25/1.25}$ or $B_m^{1.8}$, which compares in a more favorable manner than previously with the usual value of the power of B_m.

Information Theory and Its Application to Taxonomy

D. K. C. MacDonald*
Physics Division, National Research Council, Ottawa, Canada
(Received December 31, 1951)

A possible source of confusion between the concept of information content and entropy in the theory of information is discussed and resolved. Information theory is then applied to the problem of taxonomy or classification of data and several models are discussed, representative of various possible methods of filing data with the purpose of determining the optimum size of filing-unit in relation to the given data. The possible fields of application envisaged include the classification of faults in a system, the statistical analysis of data or the indexing of books, letters, etc.

I. INTRODUCTION

OF recent years, Shannon[1] (e.g., 1948, 1951), particularly, has developed a theory for the absolute measurement of the "quantity of information" contained in a system of elements capable of arrangement in various ways. If the *a priori* probability of any particular (the ith) arrangement be p_i, then the information content is defined by Shannon as

$$H=-\sum_i p_i \log p_i. \tag{1}$$

If the logarithms are evaluated to the base 2 (binary scale), then the units of H are defined as "binary digits" or "bits." If, however, as is more convenient in analysis, we use natural logarithms (to the base e) we suggest the abbreviation of "nits."† Because of the close parallel of (1) with a familiar expression for the entropy of a statistical system as defined originally by Boltzmann, the terms "entropy" and "information" are frequently regarded as synonymous. However, entropy is essentially a measure of disorder while information would appear evidently to be a measure of order, and this has led to some misunderstanding in the past. We wish here first to resolve this anomaly (which has also been discussed by Brillouin[2] (1951)), and then to present an analysis of optimum conditions for "filing" elements of data based on information theory.

II. INFORMATION AND ENTROPY

After defining the information content as in (1) above, Shannon continues to say that ". . . we shall frequently speak of quantities having the form $-\sum p_i \log p_i$ as *entropies*. . . ." On the other it appeared particularly evident at the Symposium on Information Theory (London, 1950) that many members felt confused because entropy measures a state of disorder while surely information must have the reverse quality. In particular, there was considerable discussion then about the meaning of "information content of Gaussian noise." The situation is underlined by a passage in Wiener's book[3] *The Human Use of Human Beings*—"A measure of information is a measure of order. Its negative will be a measure of disorder and will be a negative number. It can be made artificially positive by adding a constant quantity or starting from some value other than zero. This measure of disorder is known to the statistical mechanicist (sic) as *Entropy*. . . ."

It appears that this paradox is resolved if one treats entropy as an absolute measure of the over-all state of a system while information ought to be regarded as essentially an incremental quantity, characteristic of a (possible) transition of the system from one state to another. Symbolically then we might say that

$$I_{12}=-\Delta S_{12} \tag{2}$$

at least in the whole field of communication. Thus we may certainly say that the entropy of a signal source (in a quite general context) is

$$S=-\sum_i p_i \log p_i$$

for an ensemble of available, or possible, messages of relative probability p_i before, of course, any particular "message" is singled out for "transmission." It appears unwarranted, however, to call this quantity the information content without further proviso. What we can say is that if in fact any *specific* message is chosen for transmission then the information (gain) will be $-\sum_i p_i \log p_i$ (i.e., $-\Delta S\equiv-(S_2-S_1)=-0+S_1$) for then, of course, the entropy is reduced from S_1 to zero by the selection. On the other hand, if a less restrictive choice were made, —say of the group of "messages:" l, m, n, —, then the information (gain) will now be

$$S_1-(-\sum_{i=l}^{n} p_i \log p_i)$$

which is of course less than in the previous case. Similarly, when Shannon[1] defines transmission rates of channels, thus

$$R=H(x)-H_y(x) \tag{3}$$

(x being the input signal, and y the output), this ap-

* Lately of Clarendon Laboratory, Oxford.
[1] C. E. Shannon, Bell System Tech. J. **27**, 376, 623 (1948); **30**, 50 (1951).
† Despite its somewhat unattractive entomological significance.
[2] L. Brillouin, J. Appl. Phys. **22**, 334 (1951).
[3] N. Wiener, *The Human Use of Human Beings* (Houghton Mifflin Company, Boston, Massachusetts, 1951).

pears in entire agreement with this concept. First, a signal "x" having been chosen for transmission there is a *drop* of entropy $H(x)$ (to zero)—i.e., a gain of information $H(x)$—while any subsequent uncertainty in the relationship of "y" to "x" (resulting in $H_y(x)>0$) contributes a loss of information and an increase of entropy. In brief, before the transmission operation the over-all system has entropy $H(x)$ and, at the conclusion, $H_y(x)$ corresponding to a net gain of information: $H(x)-H_y(x)$.

III. TAXONOMY OR "FILING"

If we wish to classify data (letters, faults, botanical elements . . .) in a set of cells (files, classes, species . . .), it is evident that if the cell-size is made too small in relation to the "extent" or coverage of a typical element of the data then an uncertainty will enter. If an element of data is indeed found to extend over more than one cell, then some fresh course of action must be adopted. We shall consider here two possible "filing instructions:" the introduction of a "miscellaneous" cell for such elements and secondly the use of "cross-referencing"—that is, effectively duplicating the element so that it may be filed in two or more cells.

As the simplest model of this problem we consider filing an element of given finite length or extent, say l, in a line of given length, L, divided into N contiguous segments of length m, say, and inquire about the optimum length of m (choice of typical file-size) to give the greatest potential information content.

Consider then "throwing at random" an element onto the line; we may say that the probability p_i, which will fall completely inside any one arbitrary segment $m_i(=m$ in length, all m_i assumed equal and of equal *a priori* probability), is

$$p_i=(m-l)/L. \tag{4}$$

(This, of course, reduces to $p_i=m/L$ for "point" elements.) Then the total probability that it will fall inside some segment is

$$P=\sum_{i=1}^{N}\frac{m-l}{L}=N\left(\frac{m-l}{L}\right)=1-\frac{l}{m}. \tag{5}$$

Thus, the complementary probability that it will have to be placed in the (extra) "miscellaneous" segment is clearly

$$1-P\equiv q, \quad \text{say}=l/m.\ddagger \tag{6}$$

Hence, the information content (per element of data) is

$$H=-\{\sum_{i=1}^{N}p_i\log p_i+q\log q\}. \tag{7}$$

(Where logarithms are to base e, for convenience),

$$\text{i.e.,}\quad H=-\left\{N\left(\frac{m-l}{L}\right)\log\left(\frac{m-l}{L}\right)+\frac{l}{m}\log\frac{l}{m}\right\}$$

$$=-\left\{\left(1-\frac{l}{m}\right)\log\left(\frac{m-l}{L}\right)+\frac{l}{m}\log\frac{l}{m}\right\},$$

$$\therefore\quad \frac{dH}{dm}=\frac{-l}{m^2}\log\left(\frac{m-l}{L}\right)-\left(1-\frac{l}{m}\right)\left(\frac{1}{m-l}\right)$$

$$+\frac{l}{m^2}\cdot\log\frac{l}{m}+\frac{l}{m^2}$$

$$=-\frac{l}{m}\log\left(\frac{m-l}{L}\right)-1+\frac{l}{m}\log\frac{l}{m}+\frac{l}{m}.$$

(If $m\neq 0$, a trivial case) yielding a maximum for H when

$$\frac{l}{m}\left(1+\log\left\{\frac{l}{m}\left(\frac{L}{m-l}\right)\right\}\right)=1$$

i.e.,

$$m=l\left(1+\log\frac{lL}{m(m-l)}\right). \tag{8}$$

It is clear then that, generally, m is optimally of the same order of magnitude as l, but the exact solution depends (rather uncritically, as we might expect, because of the logarithmic dependence) on the choice of L or rather on the "length" of the whole filing system in relation to the basic data element (l). Thus let us take $L=e^5 l$; that is $L\approx 150l$, which means in terms of a structure liable to faults that the "typical" fault involves something less than 1 percent of the whole structure. Then writing $x=m/l$, we have

$$x=6-\log\{x(x-1)\} \tag{9a}$$

with solution $x\approx 3.7$, i.e., the basic file-segment should be about 4 times the length of a data-element.

If, however, $L\approx e^{10}\cdot l$, (i.e. $L\approx 22{,}000l$ and therefore the typical fault occupies about 1/200 percent of the whole file-system), we then have

$$x=11-\log\{x(x-1)\} \tag{9b}$$

yielding $x\approx 8.0$ in this case.

Let us now consider a model representative of cross-referencing for data-elements which are found not to lie wholly within any one filing-cell. In the first place we shall assume (as might well be the case in a structure liable to localized faults) that an element can only overlap two neighboring cells. Our analysis will now depend on the type of action we take on finding an overlap and how we interpret it.

If we were concerned with, say, a letter-filing system, a simple cross-reference in the adjacent file would simply mean that the letter had now been assigned a broader

‡ We assume throughout that L is sufficiently great that we may ignore "end-effects."

location in filing-space—more exactly $2m-l$. Thus we should now have (again neglecting "end-effects")

$$H=-\left\{\Sigma\left(\frac{m-l}{L}\right)\log\left(\frac{m-l}{L}\right)+\frac{l}{m}\log\left(\frac{2m-l}{L}\right)\right\}. \quad (10)$$

It should be observed that in the second term the multiplying factor has remained unchanged on the assumption that the *a priori* probability (l/m) of an overlap is unaltered while the amount of information resulting from the operation of cross-referencing, $\log(2m-l)/L$, is now of course different. Minimizing H with respect to m and writing $m/l=x$, as before, we find the optimal condition

$$x=\left(\frac{2x-1}{2x+1}\right)\log\left(\frac{2x-1}{x-1}\right). \quad (11)$$

This yields $x\approx 1.06$, and it is interesting to note that this is now independent of the value of L.

In the case, however, that we are concerned in our classification system with the direct accumulation of statistical information, the case of an overlap may mean that we can utilize this as locating an element in the overlap region. This then leads to

$$H=-\left\{\left(1-\frac{l}{m}\right)\log\left(\frac{m-l}{L}\right)+\frac{l}{m}\log\frac{l}{L}\right\} \quad (12)$$

and now the optimal equation is

$$x=-\log(x-1), \quad (13)$$

yielding

$$x\approx 1.28, \text{ again independent of } L.$$

It is then interesting to compare the information-gain (or entropy-reduction) per element under these various filing-methods used optimally in each case.

In the former double-filing case we have then,

$$H_{\text{opt}}=\log\frac{L}{l}+\log\left(\frac{l}{m-l}\right)-\frac{2m+l}{2m-l} \quad (14)$$

with the optimal value of m/l, i.e.,

$$H_{\text{opt}}=[\log(L/l)+.03] \text{ nits} \quad (14a)$$

while in the latter case

$$H_{\text{opt}}=[\log(L/l)+.28] \text{ nits.} \quad (15)$$

Comparison then of this latter value with the case of the "miscellaneous" file treated first shows that the double filing system is always better, as we might expect, as long as

$$\left\{\frac{m-l}{l}+\log\frac{m-l}{l}\right\}\geq -.28,$$

i.e.,

$$L/l\geq 1.11, \quad (16)$$

and this value of L is so low anyway, of course, as to render filing almost meaningless.

It is also quite straightforward to extend the double-filing analyses to the case where the "overlap" is not necessarily confined to two adjacent file-segments, and we shall not trouble to give the details here.

Finally, we may mention that there is no difficulty in considering a system where the element of data has some statistical distribution of length. Thus if the distribution density-function of l be $p(l)$, then for a "miscellaneous-file" system, for example,

$$H=-\int_0^\infty\left\{\left(1-\frac{l}{m}\right)\log\left[\left(1-\frac{l}{m}\right)\frac{m}{L}\right]+\frac{l}{m}\log\frac{l}{m}\right\}p(l)dl. \quad (17)$$

If we take now

$$p(l)=1/l'; \quad 0\leq l\leq l'$$
$$=0; \quad l>l'.$$

Then we obtain finally for maximum information content (after some reduction)

$$\frac{m}{l'}=1+\log\left(\frac{Ll'}{m(m-l')}\right)-\left(\frac{m}{l'}\right)^2\cdot\log\left(\frac{m}{m-l'}\right),$$

which for the case of $L/l'=e^5$, considered earlier, yields

$$m/l'=2.2. \quad (18)$$

ACKNOWLEDGMENT

I am grateful to Mr. J. D. North who suggested the general problem to me.

Theory of Jet Formation by Charges with Lined Conical Cavities*

EMERSON M. PUGH, R. J. EICHELBERGER, AND NORMAN ROSTOKER
Department of Physics, Carnegie Institute of Technology, Pittsburgh, Pennsylvania
(Received January 7, 1952)

An article by Birkhoff, MacDougall, Pugh, and Taylor (see reference 1) presented hydrodynamic theories of jet formation and target penetration by explosives with lined conical cavities. However, it was unable to explain satisfactorily why the jets produced are several times as long and, therefore, several times as effective as the steady-state theory predicts. It is shown here that these difficulties are overcome by assuming a variable instead of a constant collapse velocity for the walls of the conical liner. The variability in the collapse velocity produces a surprisingly large change in the process of jet formation.

THE effects produced by high explosive charges depend to a great extent upon the geometry of the charge assembly. Charges with conical cavities which are lined with a metal have probably attracted the most interest because of their usefulness in making holes through thick armor plate or producing deep penetration into other materials. The first fairly complete explanation of these phenomena appeared (in unclassified literature) in the June, 1948, issue of the *Journal of Applied Physics*.[1] By assuming that the pressures and energies involved were so great that the strength of the metals could be ignored and that, therefore, these metals could be treated as perfect fluids, the authors were able to produce mathematical theories, based upon classical hydrodynamics, that quantitatively explained most of the experiments performed upon these phenomena during the late war. They were able to show that a detonation wave sweeping from apex to base along a conical liner collapses the liner into a small diameter jet shooting forward with velocities much greater than that of the highest speed rifle bullet. They were also able to show that when this high speed jet impinged upon the target material it produced pressures close to a million atmospheres which forced the target material to flow plastically out of the path of the jet.

The mathematical theory of penetration showed that the depth of the hole produced is proportional to the length of the jet and, surprisingly, independent of the velocity of the jet, provided the velocity is great enough that the strength of the target can be neglected.

The dependence of the depth of the hole upon the length of the jet explained why deeper holes are produced when the charge with the lined conical cavity is detonated at some distance from the target. It was learned experimentally at an early date that the front end of these jets had higher velocities than the rear and that the velocities varied continuously from front to rear. This gradient in velocity along the jet causes it to lengthen as it travels. Increasing the standoff gives time and space for the jet to lengthen so that it may produce deeper holes in the target. This benefit from the increased standoff does not continue indefinitely with increasing standoff because of the tendency of the jet to spread out and become less effective.

While the theory of penetration took account of the increasing length of the jet resulting from its velocity gradient, the theory of jet formation did not explain satisfactorily why this velocity gradient exists. In fact, as recognized in the article,[2] the jets from conical liners are much longer and contain much more liner material than one should expect from the steady-state hydrodynamic theory set forth in that article. While the front of these jets followed the theory quite accurately, the rear of the jets did not. Flash radiographs of the early jet formation verified the predictions of the steady-state hydrodynamic theory. Radiographs of the later stages appeared to show jet emerging from the end of the slug long after the collapse process was completed. This gave rise to the postulate that the front of the jet was formed according to the steady-state theory but that the rear of the jet, called the "after jet," was formed by some entirely different process; possibly by extrusion or ductile drawing from the collapsed slug. The "after jet" had to account for the majority of the penetration at large standoff. If the jet were formed by the steady-state theory only, the depth of penetration would be independent of standoff and, for a conical steel liner penetrating into a steel target, would be equal only to the slant height of the cone. The penetration actually observed is from two to four times this slant height.

The present version of this theory visualizes the whole jet as formed by one single continuous process that is very much like the process previously postulated for the formation of the front part of the jet. The whole process of jet formation can be explained by adding one new assumption to the steady-state theory. According to this new assumption, the velocities with which the various elements of the cone liner collapse when they are struck by the detonation wave depends upon the original position of the element in the cone. The collapse velocities decrease continuously from apex

* This work was performed under a research contract with the Office of the Chief of Ordnance, U. S. Army, and has been released for publication by the Office of Public Information, Department of Defense.

[1] Birkhoff, MacDougall, Pugh, and Taylor, J. Appl. Phys. **19**, 563 (1948).

[2] See reference 1, p. 572.

to base.[3] The rate of decrease is very gradual near the apex but becomes much more rapid near the base. This is just what one should expect,[3] since near the base the belt of explosive surrounding the liner is much thinner and the masses of the liner elements are much greater than near the apex.

The decreasing velocity of collapse of the liner elements as the detonation wave sweeps the liner produces the remarkable effect of greatly increasing both the length of the jet and the total mass of liner material that is formed into it.

A qualitative idea of the effect of these velocity variations can be obtained from Fig. 1, if it is remembered that the velocity of the jet and the proportion of the liner that goes into the jet depend critically upon the angle β that the collapsing liner makes with the axis. As β increases the velocity of the jet decreases, but the proportion of the liner going into the jet increases. Figure 1 shows how the decreasing velocity of collapse increases the angle β.[4] While the detonation wave sweeps from P to Q along the surface of the cone APQ, the element originally at P collapses[5] from P to J. The element of cone originally at P', starting later and collapsing more slowly, arrives at M at the same time as the element from P arrives at J. If the element from P' collapsed with the same velocity as the element from P, the element from P' would reach N when the element from P reached J. Thus with constant collapse velocity the contour of the collapsing cone is conical, i.e., JNQ is a straight line. However, since the velocity of collapse from P' is less than that from P, the collapsing liner takes on the contour JMQ which is not conical. The angle β which the liner makes with the axis at J is greater than the angle β^+ which it would have made with the axis, if the collapse velocity were constant. This, of course, assumes that the liner is thin and that the velocity of each infinitesimal mass element of the liner is unaffected by its neighbors. The assumption that adjacent liner elements do not affect each other is consistent with the assumption made in the original theory that the liner acted like a perfect fluid. Elements of a perfect fluid cannot exert tensile or shearing forces on each other.

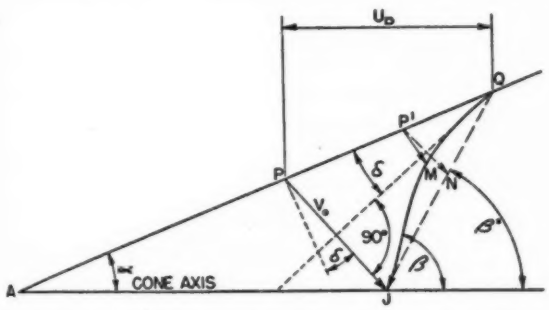


FIG. 1. Cross section showing the collapse process in a conical liner.

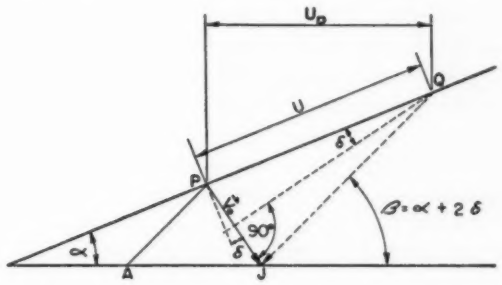


FIG. 2. Velocity vectors of an element of a collapsing conical liner.

A small decrease in the collapse velocity may produce a relatively large decrease in the jet velocity, because the jet velocity decreases both when the collapse velocity decreases and when the angle β increases. Thus a small gradient in V_0 produces a larger gradient in V_j.

THEORY OF JET FORMATION

The fact that the velocities of collapse of the cone liner decrease instead of remaining constant makes it impossible to choose coordinate systems in which true steady-state conditions exist. This is not a serious handicap, since most of the relations needed were obtained, even in the steady-state theory, by applying the laws of conservation of mass, momentum, and energy to individual zonal elements. Each element can be considered independently in the appropriate constant velocity coordinate system. Because there was some evidence indicating that the detonation velocity might slow down when it reached the thin belt of explosive around the base of the liner, accurate measurements of this velocity were made. No change in the detonation velocity U_D could be detected.

Since the velocity U_D of the detonation wave is constant, the impulse given to the liner by the detonation wave can be studied in a single inertial system. Consider a system of coordinates moving with constant velocity $U = U_D \sec\alpha$ from P to Q in Fig. 1. If the impulse per unit mass given to the liner by the detonation wave were constant, a steady-state would be set up in this coordinate system. As shown in Fig. 2, where the detonation wave is in contact with the liner, the liner would enter this region with a velocity U along the line QP and leave with the same velocity along the line PA. Since the pressures resulting from the detonation wave are everywhere perpendicular to the motion of the liner in these coordinates they can change only the direction and not the magnitude of the liner velocities. In Fig. 2, QJ is drawn parallel to PA and equal in length to QP. If the magnitudes of QP and QJ are equal to U, they represent the velocities in the moving

[3] This assumption is an obvious one and many have suggested it. The flash radiographs, however, were thought to prove that the "after jet" could not be so simply explained. However, a remarkable effect produced by the gradient in the collapse velocities had been overlooked.

[4] This increase in the angle β as the collapse proceeds towards the base of the cone is shown in the article by J. C. Clark, J. Appl. Phys. **20**, 363 (1949), Fig. 13.

[5] The details of a similar collapse process have been treated by Theodore E. Sterne in "A Note on Collapsing Cylindrical Shells," J. Appl. Phys. **21**, 73–74 (1950).

coordinate system of the liner elements entering and leaving the region around P. The vector $PJ=\overline{\mathbf{V}}_0$ is the velocity of the collapsing liner element in a stationary system of coordinates. The arguments given in the earlier paper then are valid. It is evident that the liner element does not move perpendicular to its original surface but, rather, along a line that makes a small angle δ with the normal. From trigonometry in Fig. 2, this angle is given, as in the earlier theory, by

$$\sin\delta = V_0/2U. \tag{1}$$

If the impulse per unit mass given by the detonation wave to the liner is not constant but decreases, the analysis is the same, provided it is assumed that adjacent elements of the liner do not interact with each other. The small angle δ merely becomes smaller according to the above equation. V_0 and δ are variables dependent upon the shape of the charge and the original position of the element in the cone. Referring to Fig. 1, $\delta=(\beta^+-\alpha)/2$. In the steady-state where V_0 is constant, $\beta^+=\beta$. If V_0 is considered constant and these relations are substituted for δ in the equations that follow they

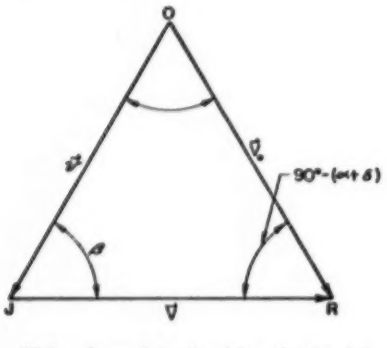


FIG. 3. Vector velocity relations at the moving junction.

will be found to be identical with the equations in the early paper.[1]

Where a ring element of the collapsing cone intersects the axis of symmetry, the analysis is more difficult because a coordinate system traveling with this junction is an accelerated system. From the analysis in the steady-state system[1] it is known that each element of mass dm splits at this junction into two parts; one part of mass dm_j goes into the jet and the other of mass dm_s goes into the slug. For each element there are four unknown quantities dm_j/dm, dm_s/dm, V_j, and V_s, where V_j and V_s are, respectively, the velocities of dm_j and dm_s in the stationary system of coordinates. From symmetry it is obvious that these velocities must be parallel to the axis. To find the four unknown quantities there are three relations available; the equations of conservation of mass, energy, and momentum. One more relation is needed. In the steady-state system this additional equation was supplied by the Bernoulli equation which cannot properly be applied to a single element and which is valid in the steady-state but not in an accelerated system.

Choose a coordinate system moving with a constant velocity V_1 equal to the instantaneous velocity of the junction when the element considered intersects the axis. At this junction the element of mass dm divides into two elements of mass dm_j and dm_s which move in opposite directions along the axis. Let these two elements have velocities with respect to the moving coordinates of v_j and v_s, respectively, where the direction of V_1 is considered positive. Then $V_j=V_1+v_j$ and $V_s=V_1+v_s$. The choice of coordinate system causes the liner element dm to move in toward the junction with a relative velocity that is parallel to the surface of the liner element. Let the absolute value of this relative velocity be v. If V_1 and β were the same for all elements, as in the steady-state system,[1] Bernoulli's theorem could be used to show that $v_s=-v$. Let us assume that $v_s=-v$ in this case also, since it is the most reasonable assumption that can be made for these quantities. It merely assumes that the liner element changes the direction but not the magnitude of its velocity as it goes past the junction in the moving coordinate system. Conservation of energy now demands that $v_j=v$; thus

$$v_j=-v_s=v. \tag{2}$$

The geometrical relations at the moving junction are shown in Fig. 3. The axis of the cone is along JR, and OJ is the element of the liner which is moving toward the axis. This element has a velocity $\mathbf{OR}=\mathbf{V}_0$ in stationary coordinates and a velocity of $\mathbf{OJ}=\mathbf{v}$ in moving coordinates. The velocity of the moving coordinates is $\mathbf{JR}=\mathbf{V}_1$. By the law of sines from Fig. 3

$$v=V_0\cos(\alpha+\delta)/\sin\beta \tag{3}$$

and

$$V_1=V_0\cos(\beta-\alpha-\delta)/\sin\beta. \tag{4}$$

In fixed coordinates the velocities of the jet and slug elements are, respectively,

$$V_j=V_1+v_j=V_1+v$$

and

$$V_s=V_1+v_s=V_1-v.$$

Substituting Eqs. (3) and (4) in these and making some well-known trigonometric simplifications, the velocities of jet and slug elements can be written

$$V_j=V_0\,(\csc\beta/2)\cos(\alpha+\delta-\beta/2) \tag{5}$$

$$V_s=V_0\,(\sec\beta/2)\sin(\alpha+\delta-\beta/2), \tag{6}$$

respectively.

In the special case, treated in the earlier article, where the velocity of liner collapse V_0 is constant, the angle β is also constant and is given by $\beta=\beta^+=\alpha+2\delta$. In this special steady-state case then, Eqs. (5) and (6) become

$$V_j=V_0\,(\csc\beta/2)\cos(\alpha/2) \tag{5a}$$

and

$$V_s=V_0\,(\sec\beta/2)\sin(\alpha/2). \tag{6b}$$

These two equations are identical with Eqs. (2) and (3) in the earlier article, though in somewhat simpler form.

Equation (1) can be used to eliminate δ from Eqs. (5) and (6)

$$V_j = V_0 (\csc\beta/2) \cos(\alpha - \beta/2 + \sin^{-1} V_0/2U) \quad (7)$$

$$V_s = V_0 (\sec\beta/2) \sin(\alpha - \beta/2 + \sin^{-1} V_0/2U). \quad (8)$$

Equations (7) and (8) are valid either in the steady-state case when V_0 is the same for all liner elements or in the nonsteady case where V_0 is different for each liner element. In the steady-state case, however, β could be expressed in terms of α, U and V_0 and need not appear in the equation.

Two of the four unknowns, dm_j/dm and dm_s/dm, remain to be calculated. As in the steady-state theory these can be calculated by combining the conservation of mass equation

$$dm = dm_j + dm_s$$

with the conservation of momentum equation

$$-dm\, v \cos\beta = dm_j\, v_j + dm_s\, v_s$$

or, since $v_j = -v_s = v$,

$$dm \cos\beta = dm_s - dm_j.$$

The resulting equations,

$$dm_j/dm = \sin^2\beta/2 \quad (9)$$

and

$$dm_s/dm = \cos^2\beta/2 \quad (10)$$

are identical with those obtained in the steady-state theory which were labeled (4) in the first paper.[1]

Equations (7), (8), (9), and (10) describe the way in which each element of the cone liner is divided into jet and slug and give the velocities of these jet and slug elements. They depend upon the cone angle 2α and the detonation velocity $U_D = U\cos\alpha$, both of which are easily measured constants; and upon the collapse velocity V_0 and the angle β, both of which are assumed to be variable. The collapse velocity V_0 depends upon the magnitude of the impulse given to the liner element by the detonation wave. No relation has yet been given for determining β. In the steady-state case previously assumed, the collapsing liner remained conical in shape but with an apex angle larger than that of the original cone. Thus the contour of the liner in the cross section shown in Fig. 1 was the straight line JQ and the angle β[+] could be determined by simple trigonometry. However, with V_0 different for each element of the liner, the contour of the collapsing liner is not straight but curved[6] like the line JMQ in Fig. 1. Thus β depends upon the shape of the collapsing liner in the immediate neighborhood of J. Let the cylindrical coordinates of M in Fig. 1 be (r, z) and the coordinates of the original position P' of M in the liner be $(x\tan\alpha, x)$. Then,

$$z = x + V_0(t-T)\sin A \quad (11)$$

and

$$r = x\tan\alpha - V_0(t-T)\cos A, \quad (12)$$

where t is the time elapsed since the detonation wave passed the apex of the cone, $T = x/U_D = x/U\cos\alpha$ and $A = \alpha + \delta$. The slope of the contour of the collapsing liner at any given time t can be obtained by differentiating r with respect to z while t is kept constant. Then from (12), since $\partial T/\partial z = (\partial T/\partial x)\partial x/\partial z = (1/U_D) \times \partial x/\partial z$,

$$\partial r/\partial z = (\partial x/\partial z)\{\tan\alpha - V_0'(t-T)\cos A + V_0\cos A/U + V_0A'(t-T)\sin A\} \quad (13)$$

where

$$V_0' = \partial V_0/\partial x \text{ and } A' = \partial A/\partial x.$$

To eliminate $\partial x/\partial z$ from (13) differentiate (11) with respect to z.

$$1 = (\partial x/\partial z)\{1 - V_0'(t-T)\sin A - V_0\sin A/U_D + V_0A'(t-T)\cos A\}. \quad (14)$$

Dividing (13) by (14)

$$\frac{\partial r}{\partial z} = \frac{\tan\alpha - V_0'(t-T)\cos A + V_0\cos A/U_D + V_0A'(t-T)\sin A}{1 + V_0'(t-T)\sin A - V_0\sin A/U_D + V_0A'(t-T)\cos A}. \quad (15)$$

which is the slope of the contour of the collapsing liner at any given time t. The time when a given element reaches the axis can now be obtained by substituting $r = 0$ in Eq. (12) which gives

$$t - T = \frac{x\tan\alpha}{V_0\cos A}. \quad (16)$$

Substituting (16) into (15) gives $(\partial r/\partial z)_{r=0} = \tan\beta$. However, before making this substitution it is desirable to make some further simplifications. Equation (1) gives $2\sin\delta = V_0/U = (V_0\cos\alpha)/U_D$. Differentiating this with respect to x gives

$$2\delta'\cos\delta = V_0'/U \text{ or } \delta' = V_0'/(2U\cos\delta),$$

and since

$$A = \alpha + \delta,\ A' = \delta' = (V_0'/2U\cos\delta) = (V_0'/V_0)\tan\delta \quad (17)$$

When (1), (16), and (17) are substituted into (15) the following simplified form is obtained.

$$\tan\beta = \frac{\sin\alpha + 2\sin\delta\cos A - x\sin\alpha(1 - \tan A\tan\delta)V_0'/V_0}{\cos\alpha - 2\sin\delta\sin A + x\sin\alpha(\tan A + \tan\delta)V_0'/V_0}.$$

[6] This curvature can be observed in the original negatives of flash radiographs of collapsing cones. See reference 3.

Since β^+ is the value of β that would be obtained if V_0' were zero, from Fig. 1, $2\delta=\beta^+-\alpha$ and $2A=\beta^++\alpha$ and the last equation can be further simplified as follows,

$$\tan\beta=\frac{\sin\beta^+-x\sin\alpha(1-\tan A\tan\delta)V_0'/V_0}{\cos\beta^++x\sin\alpha(\tan A+\tan\delta)V_0'/V_0}. \quad (18)$$

The quantities enclosed in parentheses in both the numerator and the denominator of Eq. (18) are positive unless the cone angle 2α is very much larger than is usually used. Therefore, when the collapse velocity decreases from apex to base, i.e., when V_0' is negative, the angle β is greater than β^+. From Eqs. (7) and (9) it can be seen that the increase in the angle β decreases the jet velocity and increases the proportion of the mass of the liner flowing into the jet.

It is now possible to check experimentally the validity of the assumptions made in deriving this extension of the original theory. Eqs. (1), (7), (10), and (18) have been developed from four independent concepts. The quantities U and α are easily determined constants. The quantities δ, V_0, V_j, β, and dm_s/dm are all functions of x. If all five of these variables[7] could be determined experimentally, each of these four equations could be checked independently. At this writing, only V_j and dm_s/dm have been determined accurately enough on any given charge design to provide a reliable check. This leaves three unknown variables δ, V_0, and β in the four independent equations. Equations (1) and (7) can be used to eliminate δ and V_0, leaving two independent equations, (10) and (18) in the single variable β. In an experimental paper following this one, it is shown that the two sets of values of β, determined by Eqs. (10) and (18), agree well with each other. This provides an excellent check on the theory.

It is now possible to see why the radiographs of the later stages of jet formation appeared to show jet issuing from the slug long after the collapse process was completed. This illusion is created by the fact that the last formed jet element travels at the same speed as the last formed slug element and by the fact that the velocity gradients stretch out all of the jet elements to great lengths. The last formed jet and slug elements come from liner elements near the base of the cone. In this region $V_0\ll 2U$ and $V_0\ll xV_0'$ for the charge design under consideration. Using these approximations in Eqs. (1), (7), (8), and (18), the following approximate relations are obtained: $\beta^+\cong\alpha$, $\beta\cong 90°+\alpha$, and $V_j\cong V_s\cong V_0$. Under these conditions, the last jet and slug elements that are formed travel at the same speed.

Thus the formation of the entire jet (including the "after jet") produced by charges with lined conical cavities is explained by this simple extension of the steady-state hydrodynamic theory.

The authors are indebted to Colonel C. H. M. Roberts for many valuable suggestions.

[7] Since $\beta^+=\alpha+2\delta$ and $A=\alpha+\delta$, these five are the only variables in the four equations.

Summer Courses

A special program in Infrared Spectroscopy will be given from June 16 to June 27 during the 1952 Summer Session at the Massachusetts Institute of Technology.

The program, to be offered jointly by the Institute's Spectroscopy Laboratory and Department of Chemistry, is designed for those who wish an introduction to infrared instrumentation and laboratory methods and for those interested in the use of infrared spectra in the solution of chemical problems.

The Summer Laboratory Course in Techniques and Applications of the Electron Microscope will be given again this summer from June 16 to June 28, 1952, by the Laboratory of Electron Microscopy in the Department of Engineering Physics of Cornell University. The course, under the direction of Dr. Benjamin M. Siegel, will have Dr. James Hillier of the RCA Laboratories, Princeton, New Jersey, and Dr. C. E. Hall of M.I.T., Cambridge, Massachusetts, as guest lecturers this year.

The course is designed for those research workers, institutional and industrial, who have recently entered the field of electron microscopy or who are now planning to undertake research problems involving applications of this instrument. Further inquiries should be addressed to Dr. Benjamin M. Siegel, Department of Engineering Physics, Rockefeller Hall, Cornell University, Ithaca, New York.

Experimental Verification of the Theory of Jet Formation by Charges with Lined Conical Cavities*

R. J. EICHELBERGER AND EMERSON M. PUGH
Department of Physics, Carnegie Institute of Technology, Pittsburgh, Pennsylvania
(Received January 7, 1952)

Experimental evidence is presented that verifies the assumptions and conclusions of the article, "Theory of Jet Formation by Charges with Lined Conical Cavities," by Pugh, Eichelberger, and Rostoker. A graphical method of predicting the shape of the jet and slug at every instant is presented.

THE purpose of the experimental study reported herein is to check the new form of the theory[1] of the formation of jets by charges with lined conical cavities. This theory, together with the theory of penetration[2] previously published, forms a fairly complete picture of these interesting phenomena.

In the equations derived,[1] it is assumed that when a given zonal element of the cone-liner of mass dm is struck by a plane detonation wave traveling parallel to its axis, it immediately starts to collapse towards the axis with a velocity V_0 that remains constant until the axis is reached. Upon reaching the axis this element divides into two elements of masses dm_s and dm_j, which proceed along the axis at the constant velocities V_s and V_j, respectively. Each of these constant velocities are assumed independent of the velocities of the neighboring elements. Each of the quantities mentioned are then functions of the original position of the element in the parent cone and are expressed as functions of x measured from the apex along the axis to the plane of the zonal element; i.e., dm is the element of cone mass included between two planes perpendicular to the axis at x and $x+dx$. Thus dm, dm_s, dm_j, V_0, V_s, and V_j are all functions of x. It is convenient to define m, m_s, and m_j as the integrated values of dm, dm_s, and dm_j, respectively, between the limits $x=0$ and $x=x$.

The quantities m, m_s, m_j, V_j and the detonation velocity U_D have all been determined as functions of x. $U_D(x)$ has been found to be constant. In the new form of the theory[1] there are then four independently derived equations

$$\sin\delta = V_0/2U, \tag{1}$$

$$V_j = V_0(\csc\beta/2)\cos(\alpha-\beta/2+\delta), \tag{7}$$

$$\cos^2\beta/2 = dm_s/dm, \tag{10}$$

and

$$\tan\beta = \frac{\sin\beta^+ - x\sin\alpha(1-\tan A\tan\delta)V_0'/V_0}{\cos\beta^+ + x\sin\alpha(\tan A+\tan\delta)V_0'/V_0}. \tag{18}$$

in the three unknown functions $\delta(x)$, $V_0(x)$, and $\beta(x)$. In principle, any two of these three unknowns could be eliminated to leave two independent equations in the remaining unknown. The two sets of values for the remaining unknown, obtained from the solutions of these two equations, should be the same only if the basic assumptions are sound.

Because β could be obtained rather easily from Eqs. (10) and (18), it was chosen as the remaining unknown for the purpose of verifying the theory. Since $A=\alpha+\delta$, β does not appear on the right side of either of these equations. Equation (10) gives β directly from experiment. Equation (18) is more difficult.

The values of V_0 to be substituted in Eq. (18) were obtained from Eq. (7) after eliminating δ and β with Eqs. (1) and (10), respectively.

Experiments have been performed with a number of different types of charge. However, all of the data listed in this paper were obtained in experiments with the one type illustrated in Fig. 1. The results obtained with other types of charges indicate that the conclusions drawn in this paper hold for them also. The charge shown in Fig. 1 consists of a cylinder of cast 50/50 pentolite weighing approximately ½ lb, lined with a steel cone having an apex angle of 44° ($\alpha=22°$) and a wall thickness of 0.037 in. The charge is boostered by a 30-g pellet of pressed tetryl and initiated by a No. 8 electric blasting cap.

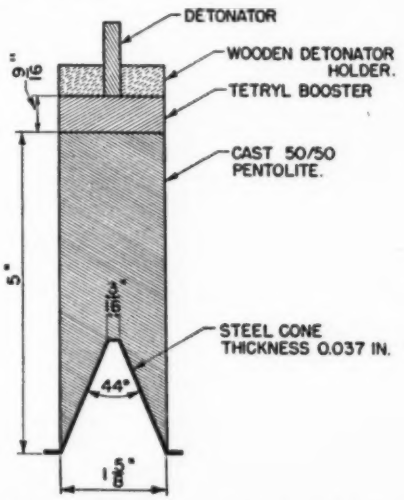


FIG. 1. Cross section of standard charge.

* This work was performed under a research contract with the Office of the Chief of Ordnance, U. S. Army, and has been released for publication by the Office of Public Information, Department of Defense.

[1] Pugh, Eichelberger, and Rostoker, J. Appl. Phys. **23**, 532 (1952).

[2] Birkhoff, MacDougall, Pugh, and Taylor, J. Appl. Phys. **19**, 563 (1948).

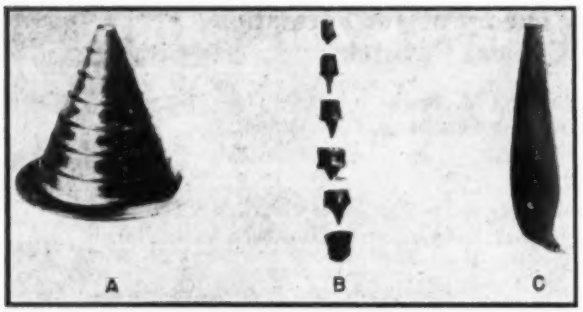


FIG. 2. Photograph of (a) grooved cone prepared for slug recovery; (b) slug sections recovered from a cone prepared as in (a); (c) a whole recovered slug from an ordinary steel cone.

The mass dm of a zonal element is determined by liner design and can be calculated with reasonable accuracy. However, to avoid errors due to lack of precision in the manufacture of the liners, samples are sectioned at accurately determined intervals and the sections weighed. This determines dm to an accuracy of better than 1 percent.

The evaluation of dm_s as a function of x was accomplished very simply. Liners were either grooved or cut into sections at regular intervals along the axis before being cast into charges. When either of these types of liner was fired into a deep container of water, the corresponding sections of the slug were recovered separately and undamaged.[3] A grooved cone is shown in Fig. 2, together with the slug sections that were recovered after firing such a liner. A whole slug recovered after firing an ordinary ungrooved liner is also shown. While the masses of the slug sections can be measured to 10^{-4} g, the statistical fluctuations limit the precision to approximately 3 percent near the apex and 0.3 percent nearer the base of the cone. Very close to the base of the liner, the mass determinations become quite erratic because a cut in this region causes the slug to break off irregularly. To minimize the influence of this effect upon the results, the average mass of a whole slug from an ordinary cone was used as an asymptote for the plot of m_s *versus* m in Fig. 3. This figure shows the results of a series of slug recovery experiments on the type of charge shown in Fig. 1. It will be observed that grooved cones yield results that are in agreement with those from presectioned cones.

The curve of m_j *versus* m in Fig. 3 was obtained from the m_s curve by subtraction, but the total mass of the jet is checked by experiments in which the whole jet is recovered by shooting it into water or ice. Neither of the quantities m_j or dm_j appear in any of the four Eqs. (1), (7), (10), or (18) that are used to check the theory. However, as will be seen later, the function $m_j(x)$ must be known to determine the function $V_j(x)$, since the experiment that measures the velocity V_j determines it as a function of m_j.

In a given explosive, the velocities of propagation of detonation waves are known to be quite constant. However, there is evidence that in very thin sticks of explosive the detonation velocity is less than in thicker ones. Since the belt of explosive around the cone becomes progressively thinner from the apex to the base, it seemed desirable to measure the detonation velocity U_D as a function of x. This has been done by several methods, but the most satisfactory was the method by S. J. Jacobs in which a photographic record is made of the travel of the luminous detonation wave by means of a rotating mirror camera. A schematic diagram of the camera is shown in Fig. 4 and a typical trace is shown in Fig. 5. This instrument uses a mirror rotating about an axis parallel to the axis of the test charge. It sweeps the image of the detonation wave along a strip of film, and thus creates a time axis normal to the direction of travel of the image of the detonation wave. To produce a sharp trace on the film, the light from the detonating

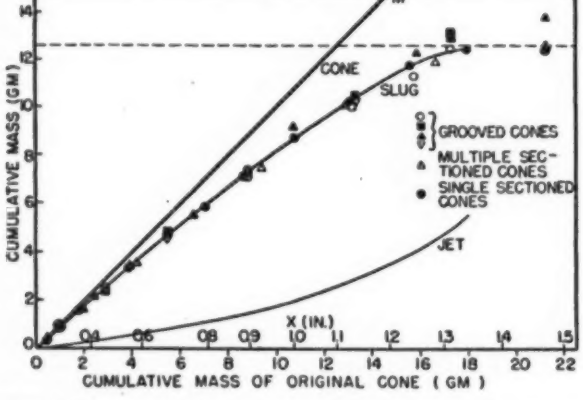


FIG. 3. Experimental correlation of the masses of slug elements with masses of parent liner elements using several types of grooved cones in addition to sectioned cones; jet mass deduced from experimental slug mass curve. Mass of whole slug is 12.6 grams.

[3] Reference 2, p. 572.

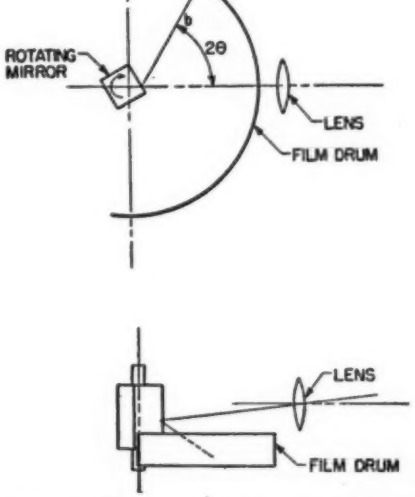


FIG. 4. Schematic diagram of the rotating mirror camera.

charge is limited by a slit placed near to the charge and parallel to its axis. The velocity is obtained from the speed of the mirror and the slope of the continuous trace. The rotational velocity of the mirror can be set at any value from 60 rps to 600 rps, and is measured to within ± 0.5 rps. The maximum writing speed in the film is 1.0 mm/microsecond. The over-all inaccuracy of these measurements was within ± 1.0 percent. While the detonation velocity measured along the side of the charge shown in Fig. 1 varies slightly with casting conditions, there is no evidence that there is a significant change in this velocity as the wave passes through the thin belt of explosive near the base of the cone. Therefore, U_D is considered independent of x, and its average value for the charges investigated was observed to be 7510 m/sec± 1.0 percent.

The velocity of the front of the jet can be measured with the rotating mirror camera just described. The jet travels so fast that when it passes through air, the front is accompanied by a luminous shock wave[4] bright enough to be recorded on the film in the same manner as detonation waves are recorded. Since each element in the jet travels with a different velocity, it is desirable to measure the velocity of all parts. This was done by methods, originated by D. P. MacDougall and his co-workers, of passing the jet through different thicknesses of target plate before the camera records its velocity. When a jet perforates a target plate, the front of it is knocked off and stays in the plate while the remaining elements pass through practically undisturbed. The velocity of the front of the remaining portion of the jet is then recorded by the camera. By using different thicknesses of target, different lengths of jet were removed and the velocity of different elements measured. A typical picture obtained by shooting a jet through a single target plate is shown in Fig. 6(a). If only the velocity of each jet element were of interest, a large number of elements of a single jet could be observed and their velocities measured by shooting the jet through a series of thin target plates with air spaces between. Such an arrangement and the resulting velocity record are shown in Fig. 6(b). This shows very graphically that the velocity of the jet elements decreases from the front to the rear of the jet. But with this multiple plate method, it is not possible to connect the velocity of the element measured with the original location of that element in the uncollapsed cone. To make this very important connection, it was necessary to shoot the jet through a single plate and collect and weigh the portion of the jet that passed through undisturbed. The difference between the mass collected and the total mass of an undisturbed jet was the m_j for the particular element whose velocity V_j was recorded. By using different thicknesses of target plate the velocity of different elements in the jet were

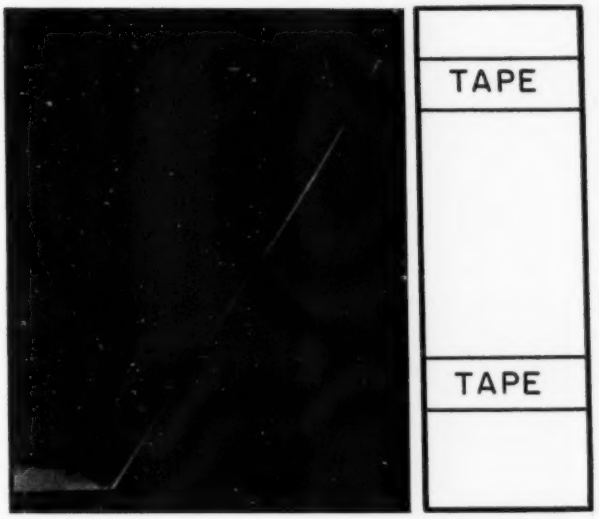


FIG. 5. Experimental set-up for detonation rate measurements and typical trace.

determined and the element measured was identified. This gave V_j as a function of m_j. Since, by means of the plot in Fig. 3, $m_j(x)$ could be obtained, $V_j(x)$ could be obtained also. A graph of V_j *versus* the mass of residual jet collected is given in Fig. 7 and the curve of V_j *versus* x is given in Fig. 8. In principle, the procedure is simple but collection of the jet presents formidable difficulties on account of the penetrating ability of the jet, and the tremendous radial pressure developed by it in any target medium. At first, a complex and unwieldy pressure resisting chamber was developed to use water as the collecting medium. In more recent experiments parts of the jet have been collected more conveniently and simply in cylinders of ice.[5] In either case, several gallons of water were employed in collecting a jet weighing about six grams and consisting of particles from 0.01 to 1.0 mm in size. This water was very carefully filtered and the residue subjected to quantitative chemical analysis to determine the mass of jet material collected. Where target plates were used to remove the front of the jet they were, of course, of different material than the jet. Considerable care was required throughout to prevent loss of jet material or contamination of the sample. The experimental error of the jet collection experiments could not be accurately evaluated, but probably amounted to about 4 percent.

According to Eq. (10) the slope of the line m_s *versus* m, in Fig. 3 gives the $\cos^2\beta/2$. Values of β obtained from this relation are plotted against x in Fig. 9. It is difficult to obtain accurate results when an experimentally determined curve must be differentiated. To gain all of the accuracy possible from the data, a polynomial up to the fifth degree was fitted to the data plotted in

[4] Pugh, Heine-Geldern, Foner, and Mutschler, J. Appl. Phys. 22, 488 (1948), Figs. 2 and 3.

[5] Similar methods of collecting jet material have been developed by the Explosives Research Laboratory under D. MacDougall, and duPont's Eastern Laboratory under C. O. Davis.

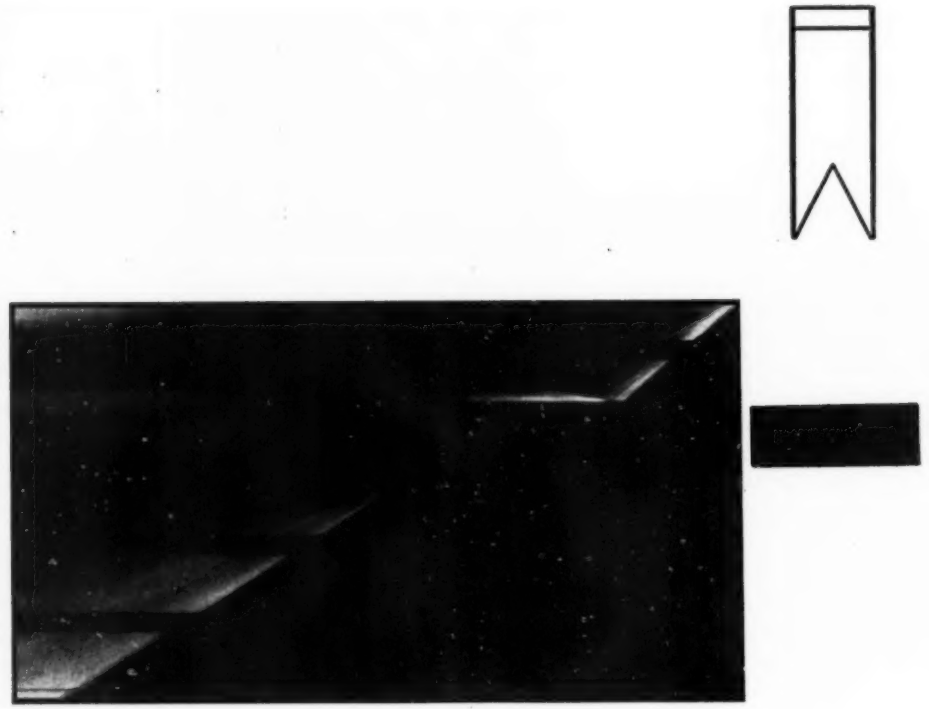


FIG. 6(a). Experimental set-up for jet velocity measurements and typical trace. Horizontal black streaks are produced by calibration tapes.

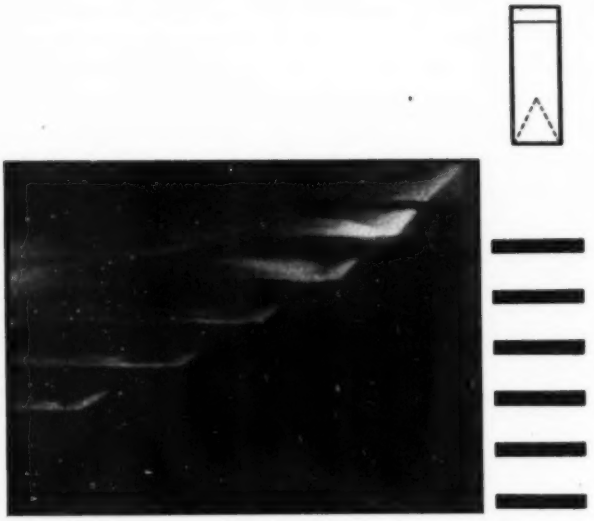


FIG. 6(b). Experimental set-up for multiple jet velocity measurements and typical trace.

Fig. 3 by the method of least squares,[6] and the values of β were determined by direct differentiation.

With the values of V_j and β available in Figs. 8 and 9, it was a simple matter to determine $V_0(x)$ from Eq. (7). The results are plotted in Fig. 10. These values of V_0 could have been substituted into Eq. (18) to obtain values of β to compare with those in Fig. 9. It was simpler, however, to substitute values from Fig. 10 into the parametric Eqs. (11) and (12) from which Eq. (18) was derived, and to plot the resulting values of z and r for various instants of time. By drawing separate curves through the points determined for each instant of time, a series of contours of the collapsing liner was obtained. The angle between a contour and the axis, where the contour intersects the axis, was the value of β for the particular element that intersects the axis. The process is illustrated in Fig. 11 where the finite dimensions of the liner are also taken into account. In this figure, an element from near the apex of the cone is shown in its original position 01 and in a later position where it is divided between jet, 0′1′, and slug 0″1″. An element some distance from the apex is also shown (cross-hatched) in its original position, and in a later position where it is collapsing toward the axis. In Fig. 12, the graphically determined values of β are plotted

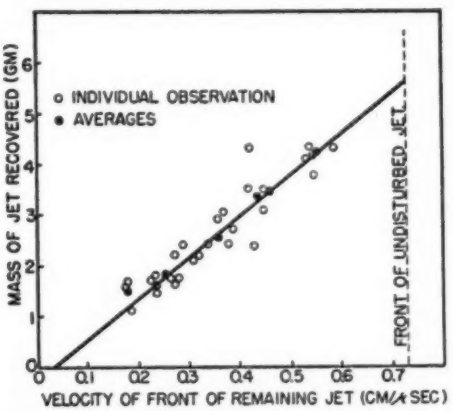


FIG. 7. Experimental correlation of mass of the unused part of the jet and the velocity of its front.

[6] Raymond T. Birge, Revs. Modern Phys. 19, 298–360 (1947).

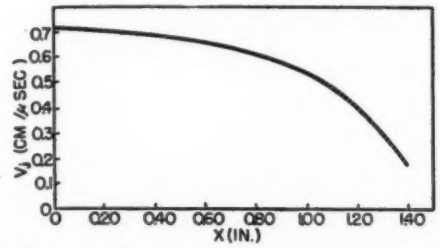


FIG. 8. Calculated correlation of the velocities of jet elements V_j with initial position x of parent cone elements.

as circles to compare them with the values determined by slug recovery experiments, which are represented by a solid line. The agreement is better than could be expected from the accuracy of the experiments.

It is easy to see how the illusion of the "after jet" arose; for the velocity of collapse decreases very gradually, at first, (simulating steady-state conditions) near the apex but decreases much more rapidly as the wave approaches the base. This causes the X-ray pictures to look like jet material is issuing from the slug long after the collapse process appears to be completed.

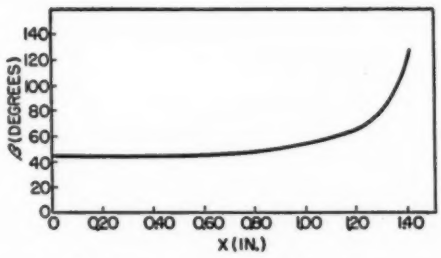


FIG. 9. Calculated angle of collapse β as a function of the initial position x of the parent cone elements.

However, there is no indication of sudden change in the experimental data that can justify dividing the phenomenon into two parts. In fact, it is a strange coincidence that, with the charge investigated, the curves of m, β, and V_0 shown in Figs. 3, 9, and 10, which are all nonlinear with x, combine together to produce in the jet a uniform distribution of mass and a linear distribution of velocity. The fact that the zonal elements are much more massive near the base of the cone, combined with the fact that β increases rapidly in this region, leads one to expect that the rear of the jet should contain much more mass than the front and should therefore be much larger in diameter; but the velocity

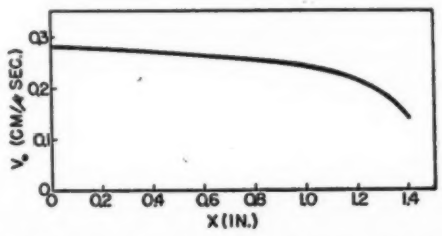


FIG. 10. Calculated velocity of wall collapse V_0 as a function of the initial position x of the parent cone elements.

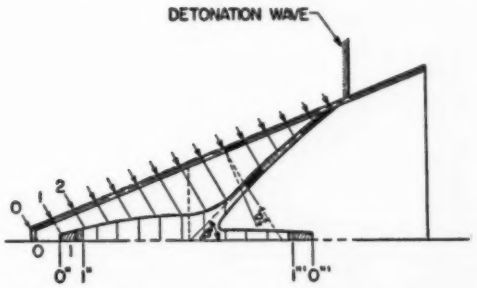


FIG. 11. Illustration of method of graphical solution. The positions of liner elements 0, 1, 2,—or the slug and jet elements produced by them are calculated for any arbitrary instant of time. The magnitude of V_0 and the angle δ are calculated from equations $V_0 = V_j \ (\sin\beta/2)/[\cos(\alpha+\delta-\beta/2)]$, and $\delta = \sin^{-1}\{V_0 \cos\alpha/2U_D\}$. The position of an element such as 9 or 10 is then determined by the product of V_0 and the time elapsed since the detonation wave passed the element. The conservation of mass is sufficient to determine the thickness of any element of the collapsing wall. For an element such as 0 or 1, which has already reached the axis and split up into jet and slug elements at the instant under consideration, the time since they reached the axis at 0′ and 1′, respectively, is obtainable; the velocity of the resulting jet element is known and that of the slug element can be calculated; thus the positions 0″ and 1″ of the slug elements, and 0‴ and 1‴ of the jet elements, are determinable. The distribution of the liner mass between jet and slug is obtained from equation $\beta = 2\cos^{-1}(dm_s/dm)^{\frac{1}{2}}$, and serves to calculate the jet and slug diameters. The collapse angle β is the angle between the axis and a tangent at the axis to the central "streamline" of the collapsing wall.

of collapse of the walls decreases just rapidly enough in this region to keep the rate at which mass goes into the jet uniform. The uniform distribution of mass is shown in Fig. 7. The linear velocity gradient is shown in Fig. 13.

The method illustrated in Fig. 11 can be extended to show, step by step, how the jet is formed as the cone collapses. The results of such analysis are shown in Fig. 14.

These results can be compared with some flash radiographs by Clark.[7] Although his charges were of different design than the one shown in Fig. 4, the

FIG. 12. Comparison of results of graphical solution of parametric equations with values of β from slug recovery experiments.

[7] J. C. Clark, J. Appl. Phys. **20**, 369 (1949), Fig. 13.

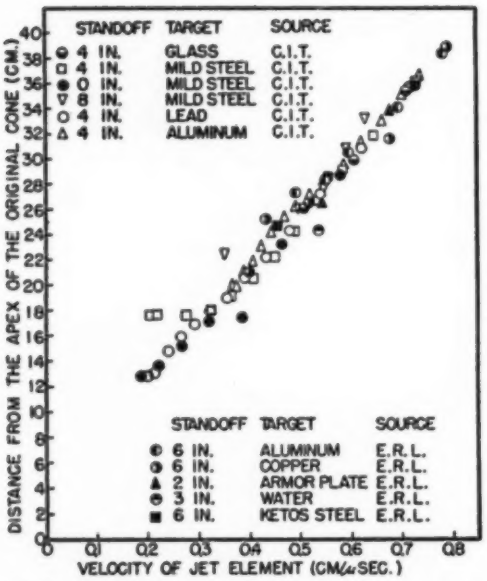

STANDOFF TARGET SOURCE
4 IN. GLASS C.I.T.
4 IN. MILD STEEL C.I.T.
0 IN. MILD STEEL C.I.T.
8 IN. MILD STEEL C.I.T.
4 IN. LEAD C.I.T.
4 IN. ALUMINUM C.I.T.
STANDOFF TARGET SOURCE
6 IN. ALUMINUM E.R.L.
6 IN. COPPER E.R.L.
2 IN. ARMOR PLATE E.R.L.
3 IN. WATER E.R.L.
6 IN. KETOS STEEL E.R.L.
DISTANCE FROM THE APEX OF THE ORIGINAL CONE (CM.)
VELOCITY OF JET ELEMENT (CM/μ SEC.)

FIG. 13. Location of jet elements 50μ sec after the detonation wave passed the apex of the cone, plotted as functions of the velocities of the elements. A wide variety of target plates placed at different standoffs were used in obtaining this data.

similarity between the predicted shapes shown in Fig. 14 and the radiographs is striking.

The curve plotted in Fig. 13 was obtained from some additional experimental data that was not needed in the test of the theory just described. Since this data was obtained with targets made of several different kinds and thicknesses of material placed at several different standoffs, it provides experimental justification for the assumption that, during target penetration, adjacent elements of the jet do not interact appreciably with each other. The data was obtained from experiments with a rotating mirror camera like the one that produced the trace shown in Fig. 6(a). In Fig. 6(a) no use has yet been made of the fact that the horizontal distance between the point where the trace disappears into the target and the point where it emerges again measures the time required to perforate the target plate. If it is assumed that the element making the emergent trace (because it is now at the front of the remaining jet) has not had its velocity changed by the process of penetration, one can calculate the position of this element at any desired time. By varying the thickness of the target plate, one can select different elements having different velocities and calculate their locations at any desired time. Assuming, then, that jet elements do not change their velocity during the time measurements are made, one can calculate the positions of each of the elements at any given time. Figure 13 shows the location of each jet element 50 μsec after the front of the jet passed the base of the original cone. The fact that the results are the same, with the various different standoffs, indicates that the velocities of jet elements remain constant in flight, and the fact that the results are the same with different target materials indicates that adjacent elements do not affect each other appreciably during penetration. Obviously these assumptions cannot be strictly true. The experimental evidence only shows that neither the frictional resistance of the air nor the plastic or elastic stresses within the jet are great enough to noticeably affect the velocities of the jet elements. By the time one of these jets travels forty feet or more, the frictional drag of the air becomes quite great, largely because by that time the jet has become widely separated into discrete particles. As long as the particles are in close-packed jet formation, the air resistance has a very small effect upon their velocities. It has not been possible to detect any change in the velocity of fronts of jets in their first two feet of travel; i.e., no curvature has been observed in traces made by the rotating mirror camera of jets traveling this distance.

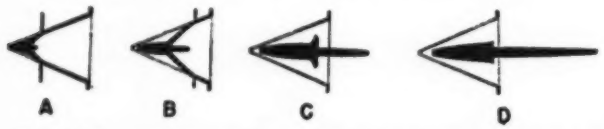


FIG. 14. Stages of cone collapse for an *S*-1 charge, calculated for (a) 2½μ sec, (b) 5μ sec, (c) 10μ sec, and (d) 15μ sec intervals after the detonation wave passed the apex of the cone.

CONCLUSIONS

The experimental evidence furnishes strong proof that jets from charges with lined conical cavities are produced in one continuous process and that this process is very similar to the hydrodynamic process previously proposed to explain only the fronts of these jets. The restrictive assumption in the earlier theory, that all of the processes were steady in properly chosen constant-velocity coordinate systems, made it impossible to explain adequately the formation of more than the first part of the jet and made it necessary to postulate some other process for the formation of the last part which was called the "after jet." It is interesting to notice that the new form of the theory naturally introduces the gradient of velocity along the jet. This was known to exist when the older paper was published, but its formation was not adequately explained by the older theory. Presumably one could design charges in which the collapse velocity would remain constant in conformity with the assumptions of the older form of the theory, but this would eliminate the velocity gradient that lengthens the jet and thus reduce the penetration to less than one-fifth of that now achieved. The experiments indicate that the design shown in Fig. 4 is very fortunate in that it produces just the right gradient in collapse velocity to provide in the final jet both a uniform distribution of mass and a linear distribution of velocities.

Several of the techniques employed in this test of the theory were originally introduced during World War II by members of the Explosives Research Laboratory, Bruceton, Pennsylvania. The authors feel especially indebted to G. B. Kistiakowsky, D. P. MacDougall, G. H. Messerly, M. A. Paul, and S. J. Jacobs.

The Condition for the Creeping of a Liquid Past a Mercury Seal

R. S. Jessup
National Bureau of Standards, Washington, D. C.
(Received December 21, 1951)

An observation of the creeping of a liquid past a mercury seal is reported. The conditions under which such a phenomenon might be expected to occur are discussed in terms of the thermodynamics of surfaces.

A BUNSEN-TYPE calorimeter has been constructed at the National Bureau of Standards, using diphenyl ether (melting point about 27°C) instead of ice. The original design of the calorimeter was such that changes in volume as a result of partial melting or freezing of the diphenyl ether caused mercury to flow out of or into the calorimeter through a stainless steel tube (about 2 mm in outside diameter) the lower end of which was immersed to a depth of 2 cm or more below the surface of a pool of mercury in the bottom of the calorimeter. It was found that the mercury seal thus formed was not effective in that diphenyl ether would creep past it and flow out of the calorimeter through the stainless steel tube at a rate estimated to be in the neighborhood of 0.01 ml per day. The same phenomenon has been observed, both with diphenyl ether and with tertiary butyl alcohol, in apparatus specially constructed to simulate conditions in the calorimeter. The phenomenon was not observed over a period of about a year with either toluene or diphenyl ether in similar all-glass apparatus. The following discussion indicates that the effect can be interpreted semiquantitatively in terms of the thermodynamics of surfaces.

Consider a system, such as that illustrated in Fig. 1, which is maintained at constant temperature and pressure. The initial state of the system is represented by (a) with mercury m in contact with the solid, s, and with the liquid d above the mercury. Suppose that a film of the liquid d creeps down between the mercury and the solid until an equilibrium state represented by (b), Fig. 1, is reached. Since this change takes place spontaneously, it must be accompanied by a decrease in the free energy, $F=U-TS+pV$, of the system. Considering only a vertical strip of the film 1 cm wide, that part of the change in free energy of the system in going from state (a) to state (b), which is associated with an elementary area $\Delta a(=\Delta h)$, is given approximately by,[1a, 1b, 2]

$$\Delta(F_b-F_a)=[(\rho_m-\rho_d)ght+(\gamma_{dm}+\sigma_{ms}-\sigma_{ds})]\Delta h. \quad (1)$$

Here ρ_m and ρ_d are the densities of mercury and the liquid d, respectively; g is the acceleration due to gravity; h is the vertical distance from the horizontal $d-m$ interface to the element of area, Δh; t is the thickness of the film at this element; γ_{dm} is the interfacial tension between d and m; and σ_{ms} and σ_{ds} are the adhesion tensions between the solid s and the liquids m and d, respectively.

If we denote $(\sigma_{ds}-\sigma_{ms}-\gamma_{dm})$ by Γ, Eq. (1) may be written

$$\frac{\Delta(F_b-F_a)}{\Delta h}=(\rho_m-\rho_d)ght-\Gamma. \quad (2)$$

The total change in free energy of the system (per centimeter width of film) in going from state (a) to state (b) will be obtained by integrating Eq. (2) with respect to h, so that the free energy of the system in the final state (b) will be given by

$$F_b=F_a+\int[(\rho_m-\rho_d)ght-\Gamma]dh, \quad (3)$$

where the integration is performed over all values of h for which the film exists. Since $F_b<F_a$, the integral must be negative, and since the film has spread spontaneously over each of the elementary areas Δh, it follows that the integrand must be negative over the whole range of integration or that

$$(\rho_m-\rho_d)ght<\Gamma \quad (4)$$

in all parts of the film. Since the left-hand side of the inequality (4) is positive, we must also have as a necessary condition for the formation of the film

$$\Gamma>0. \quad (5)$$

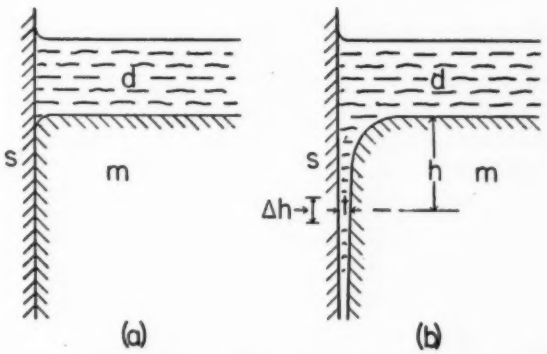


FIG. 1. Diagram representing an initial state (a) with mercury m, in contact with solid s and with liquid d, above mercury, and a final state (b) with solid and mercury separated by an equilibrium film of liquid d.

[1] N. K. Adam, *The Physics and Chemistry of Surfaces* (Oxford University Press, New York, 1941); (a) p. 178; (b) p. 179, Eq. (2); (c) p. 166 and the references there cited; (d) p. 186.
[2] G. N. Lewis and M. Randall, *Thermodynamics* (McGraw-Hill Book Company, Inc., New York, 1923), p. 242.

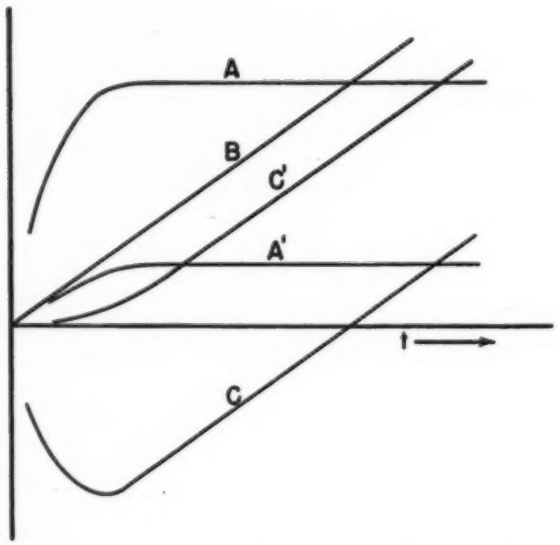


FIG. 2. Schematic diagram showing free energy relations in the formation of a film of liquid between mercury and a solid. A and A' represent two possible cases of dependence of Γ on t; B represents the function $(\rho_m-\rho_d)ght$ for some particular value of h; and C and C' represent the value of $\Delta(F_b-F_a)/\Delta h$ corresponding to A and A', respectively.

The condition that the film be in stable equilibrium is that the free energy must be a minimum. It will be seen from Eq. (3) that in order for this condition to be fulfilled it is necessary that

$$\delta\int_{h_1}^{h_2}[(\rho_m-\rho_d)ght-\Gamma]dh=0 \qquad (6)$$

for any arbitrary infinitesimal variation of t over the range between any two arbitrary values of h, say h_1 and h_2. By application of the ordinary methods of the calculus of variations it is found that, except for certain terms which appear to be negligible, the condition (6) leads to the relation

$$(\rho_m-\rho_d)gh=\partial\Gamma/\partial t, \qquad (7)$$

which must hold at the equilibrium value of t for any value of h.

According to Eq. (7) $\partial\Gamma/\partial t$ must be positive at the equilibrium value of t, and it would be expected, therefore, that at this value of t the value of Γ would be somewhat lower than that calculated from experimental values of the surface parameters, which would correspond to a thick layer of the liquid. The situation may be illustrated graphically by the schematic diagram shown in Fig. 2, where the terms of Eq. (2) for some particular value of h are plotted against t. The form of the Γ *versus* t curve A was chosen so as to illustrate a case where the spontaneous formation of a film between mercury and a solid would be thermodynamically possible. It will be seen that the minimum in the curve C representing $\Delta(F_b-F_a)/\Delta h$ is located at the value of t for which the slope of curve A is equal to that of the straight line B, i.e., to $(\rho_m-\rho_d)gh$, in agreement with Eq. (7). The form of the Γ *versus* t curve A' was chosen to illustrate a case where the spontaneous formation of a film would not be possible. In this case $\Delta(F_b-F_a)/\Delta h$, represented by curve C', is everywhere positive, and this curve has no minimum because the slope of the curve A' is always smaller than that of the straight line B.

In the absence of any data on the thickness of the film or the manner in which Γ depends upon thickness, the relations given previously cannot be applied to any particular system. However, some qualitative conclusions can be drawn from experimental values of Γ corresponding to large values of t.

Experimental (or estimated) values of Γ are given in Table I for several liquids and two solids. The surface tension of diphenyl ether was estimated by means of Sugden's parachor[1c] to be about 39 dynes per cm. The interfacial tension of mercury and diphenyl ether was taken as 375 dynes per cm, since the interfacial tensions of many organic liquids and mercury are of the order of 100 dynes per cm lower than the surface tension of mercury. The other surface and interfacial tensions were taken from International Critical Tables. The contact angles of diphenyl ether-stainless steel, diphenyl ether-glass, and water-stainless steel were estimated from rough observations to be of the order of 15°, 70°, and 90°, respectively. The contact angles of mercury-glass and mercury-stainless steel were taken as 140° and 154°, respectively.[1d] All other contact angles are given in International Critical Tables as zero or were assumed to be zero.

Although the values of Γ given for diphenyl ether and glass is practically zero, all of the values in Table I are positive, and large enough to insure that the inequality (4) would be satisfied for relatively large values of t, if the experimental values of Γ were used in this relation. As already indicated, however, these experimental values are larger than those which would be applicable. At any rate it will be seen that the values of Γ in the table are smaller for the two systems (diphenyl ether-glass and toluene-glass) in which the creeping past the mercury seal was not observed, than for the system (diphenyl ether-stainless steel) in which this phenomenon was observed.

Although the inequality (5), using an experimental value of Γ, is not sufficient to insure that the formation of a film is thermodynamically possible, the condition

$$\Gamma<0$$

TABLE I. Experimental or estimated values of Γ.

Liquid	s = stainless steel dynes/cm	s = glass dynes/cm
Diphenyl ether	78	2
Toluene	85	36
Aniline	117	66
Ethanol	73	22
Water	40	62

is sufficient to insure that a film will not be formed spontaneously. In this connection it is of interest to consider in detail the calculation of Γ for the system mercury-diphenyl ether-stainless steel. The data given above yield the values in dynes per cm, $\sigma_{ds}=38$, $\sigma_{ms}=-415$, and $\gamma_{dm}=375$. We have then for this system

$$\Gamma=(\sigma_{ds}-\sigma_{ms}-\gamma_{dm})$$
$$=(38+415-375)=78 \text{ dynes per cm.}$$

It will be seen that the large negative value of σ_{ms} is to a considerable extent responsible for the positive sign of Γ. It would be expected therefore, that creeping of the diphenyl ether past the mercury seal in the calorimeter would be less likely if a solid having a smaller contact angle with mercury were substituted for stainless steel. This was found to be the case. The lower end of the stainless steel tube was replaced by a platinum[3] tube which was plated with a thin layer of gold and then heated sufficiently to cause incipient alloying of the two metals. It was found that platinum so treated was wet by mercury, i.e., the contact angle was less than 90°, making the value of Γ for this system negative. No further creeping of the diphenyl ether past the mercury seal was observed after this change was made.

[3] The replacement of the stainless steel tube by a platinum tube was suggested by D. C. Ginnings of the National Bureau of Standards.

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 5 MAY, 1952

Electromagnetic Levitation of Solid and Molten Metals

E. C. Okress and D. M. Wroughton
Westinghouse Electric Corporation, Bloomfield, New Jersey

G. Comenetz, P. H. Brace, and J. C. R. Kelly
Westinghouse Electric Corporation, East Pittsburgh, Pennsylvania

(Received January 24, 1952)

The subject is an unconventional method of heating and melting metals without a crucible, by suspension in space with an electromagnetic field. Operating conditions for certain cases are given. The results obtained by means of the new technique encourage the thought of melting, purifying, alloying, and agitating of inert and reactive metals without resort to crucibles, and thereby avoiding the contamination of reactive metals by crucible materials. Preliminary results with various forms and masses of metal are described. Considerations concerning the atmosphere in which levitation occurs are included.

INTRODUCTION

PREVAILING techniques for melting metals involve crucibles of various types depending upon the metals. However, crucibles are often undesirable and sometimes even prohibitive. In melting the more reactive high melting point metals such as Ti, Zr, V, Ta, Mo, and others of similar chemical and physical characteristics, difficulty is usually encountered in the selection of crucible materials. When such metals are melted they react with the material of the crucible to some extent, or other reactions occur which adversely affect the quality of the melt. For example, Ti when melted in a graphite crucible may contain several tenths to 1 percent or more of carbon. When Zr is melted in contact with BeO, it becomes embrittled because of absorption of oxygen from the refractory. If the crucible is sufficiently cooled in an effort to reduce contamination of the melt then other unsatisfactory results are obtained which may or may not be tolerable.

As a consequence, a technique which reduces or prevents contamination, provides continuous agitation of the melt, and is reasonably practical was sought. In an attempt to realize this objective the following plausible methods were considered:

1. Magnetostatic levitation with simultaneous induction melting.
2. Electromagnetic or induction levitation with simultaneous induction melting.
3. Acoustic levitation with simultaneous acoustic or induction melting.

Of these methods the second appeared most promising and therefore was investigated.

Initial experiments, in air, involved a single coil. Although a brass specimen was supported by the alternating magnetic field established by the energized coil, it drifted laterally from the coil axis and fell out of the influence of the field. Force acting to maintain the specimen at the axis and above the coil was therefore inadequate. With more power the levitated specimen rotated, melted, and thereupon uncontrollably flowed down out of the influence of the coil field. The experiment was repeated with other metals with the same undesired results. However, initial experiments with brass, Sn, Al, and Ti specimens up to 550 grams indicated that in the solid state, at least, vertical support by the alternating magnetic field could be provided with the available alternator. It remained to investigate

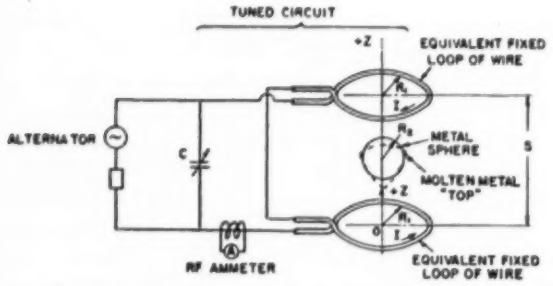


FIG. 1. Schematic diagram of the electrical system used to induction heat, melt and simultaneously levitate the solid conductive sphere.

the magnetic field distribution required for stable levitation of molten as well as solid metals, including controlled draining of the melt.

By a combination of theoretical considerations and experimental[1] work it was discovered that one way in which an adequate radial restoring force could be realized was by two coaxial coils, as illustrated in principle by Fig. 1. They are connected in series across a common capacitor so as to constitute a parallel tuned load to the alternator. The alternating currents in the (generally unequal) coils are in series opposition at all times. The currents in these fixed coils induce in the interposed metal charge, which may for example be a sphere as in Fig. 1, eddy currents whose interaction with the inducing field gives rise to the necessary supporting and lateral restoring forces. These make possible stable levitation of solid metal specimens, and are at least the principal factors in levitation of the subsequently molten metal when that can be done. Heating and melting of the levitated specimen depends upon the power available, coupling between primary circuit and specimen, frequency, eddy current, and hysteresis losses, and heat losses from the specimen.

The initial phase of the investigation took place at Bloomfield. The Bloomfield ideas and findings were communicated to East Pittsburgh, where a subsequent phase was then undertaken. In the subsequent phase, additional theoretical work and experiments in air and in vacuum were done.

THEORETICAL CONSIDERATIONS

The levitation forces on a homogeneous metal sphere of resistivity ρ in a magnetic field B produced by alternating current I in a single loop, or in two equal coaxial loops of wire as illustrated by Fig. 1, may be derived from the usual classical considerations. By such considerations the field due to the eddy currents in the relatively small sphere may be considered as equivalent to that produced by either an equatorial current I_2 or an alternating magnetic dipole. By superposition, the equivalent equatorial current can be evaluated from the magnetic field of the primary loops at the center of the sphere.[2] To illustrate the principle of levitation, the evaluation of the equivalent equatorial current is not essential.

Consider magnetostatic fields for any instant in the cycle of alternating current in the loop(s) as illustrated by Fig. 1. Then the application of the former of these two equivalent concepts leads to the following results. Essentially, the procedure involves determining the mutual inductance and hence the force between two coaxial current loops. This evaluation is made by imaging the fixed primary current loop A of Fig. 2 into the middle or equivalent secondary current loop C and then repeating the procedure, taking into account the current directions, for the primary loop B and secondary loop C. By superposition, these two forces are combined at the secondary loop to give the resultant force on this loop and hence on the sphere. Although the calculation of the forces between circuits[3] can become rather involved, for illustrative purposes the following procedure is adequate.

The electromagnetic force exerted by one circuit on another may be expressed as

$$F=\frac{1}{2}I_1^2\frac{dL_1}{dz}+I_1I_2\frac{dM_{12}}{dz}+\frac{1}{2}I_2^2\frac{dL_2}{dz}, \tag{1}$$

where I denotes the current, L the self-inductance, and M the coefficient of mutual inductance. The form and relative position of the circuits depend on the variable Z. For a rigid system

$$F=I_1I_2dM/dZ. \tag{2}$$

For the case of two parallel coaxial current loops, A

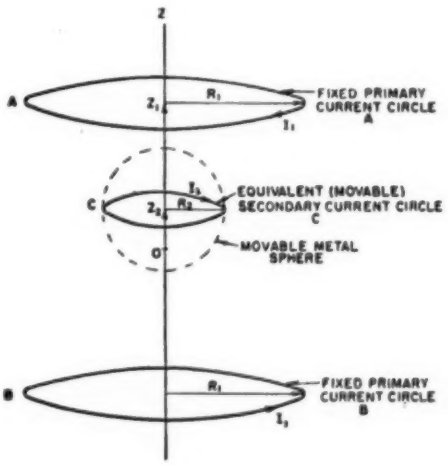


FIG. 2. Equivalent currents for fixed loops and sphere of Fig. 1.

[1] Westinghouse Research Report BL-R-940-11-1 and Research Memo 60-94472-2-5.

[2] W. R. Smythe, *Static and Dynamic Electricity* (McGraw-Hill Book Company, Inc., New York, 1950), section 11.06. In the 1939 edition, Section 11.05.

[3] See reference 2, section 7.19 of 1950 edition, or section 7.111 of 1939 edition.

and C of Fig. 2, the force between them becomes

$$F=\frac{I_1I_2\pi^2R_1^2R_2^2}{(R_1^2+Z_1^2)^2}\left[\frac{6Z_1}{(R_1^2+Z_1^2)^{1/2}}+\frac{24Z_2(Z_1^2-R_1^2/4)}{(R_1^2+Z_1^2)^{3/2}}\right.$$
$$\left.+\frac{60Z_1(Z_1^2-3R_1^2/4)(Z_2^2-R_2^2/4)}{(R_1^2+Z_1^2)^{5/2}}+\cdots\right]. \quad (3)$$

For the case of three parallel coaxial current loops, A, B, and C of Fig. 2, the expression for the force exerted on the movable equivalent current loop C by *one* of the fixed primary current loops A or B reduces to

$$F=\frac{I_1I_2\pi^2R_1^2R_2^2}{(R_1^2+Z_1^2)^2}\left[\frac{6Z_1}{(R_1^2+Z_1^2)^{1/2}}\right.$$
$$\left.+\frac{60Z_1(Z_1^2-3R_1^2/4)(Z_2^2-R_2^2/4)}{(R_1^2+Z_1^2)^{5/2}}+\cdots\right], \quad (4)$$

as the odd Z_2 terms vanish. During a time interval in which the current directions in the primary loop A and secondary loop C are the same, the force F acts to attract the movable secondary to the primary so as to increase their mutual inductance. The converse situation exists for the primary loop B and secondary loop C.

The levitation of the secondary loop is achieved by a balance between the gravitational force on the secondary loop and the longitudinal electromagnetic force due to the primary loops. The action of the transverse centrally directed restoring forces on the secondary loop is dependent upon the character and magnitude of the mutual inductance for the noncoaxial case in contrast to the coaxial case.

In the actual alternating current case, the induced currents in the metal sphere due to the fields of the two primary loops cause a force to act on the sphere in opposition to the gravitational force, when the sphere is above the lower loop but below the midplane between the loops, and hence levitation may be realized.

Presently, the explicit manner in which the restoring forces act on the sphere, to maintain it at an equilibrium position within the projected volume between the two primary fixed loops, will be illustrated graphically and with experimental confirmation.

An approximate method of obtaining a simple levitation formula in the alternating current case was conceived by the late R. C. Mason of East Pittsburgh and an independent equivalent attack was pursued at Bloomfield. The formula is again one that expresses levitation force acting upon a homogeneous sphere. If such a sphere is set into a magnetic field that is uniform in space and sinusoidally alternating in time at not too high a frequency, then the magnetic field of the eddy currents induced in the sphere is identical with that of a magnetic dipole, alternating in time with the generating field, at a phase and of a strength that depend on the generating field strength and frequency,

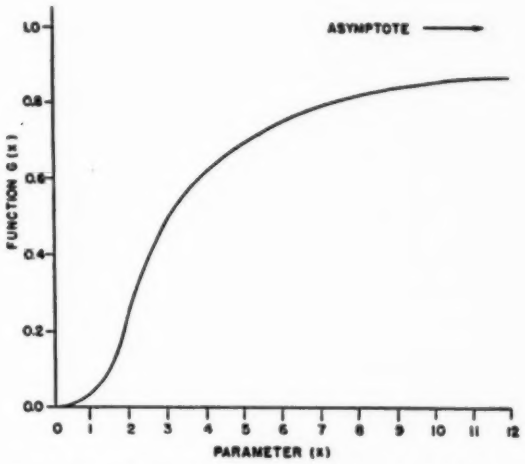


FIG. 3. The dimensionless function $G(x)$ plotted for use in formulas (5) and (6).

and on the sphere's radius, permeability, and resistivity.[4]

The net mechanical force which the uniform alternating field exerts upon the dipole, to which the sphere is equivalent, is zero. If now the sphere is set into a *non*-uniform alternating field, then to a first approximation the force may be calculated by replacing the sphere by the dipole that would be produced by a fictitious *uniform* alternating field having everywhere the value the given non-uniform field has at the center of the sphere. The actual non-uniform field would exert a force upon this dipole.

If the reader is not convinced of the validity of this approximation, he may accept it as merely plausible, and then weigh the experimental confirmation to be advanced presently.

The method leads, after some computation, to the following approximate formula for the force upon a homogeneous sphere of unit permeability in the non-uniform magnetic field generated by an alternating current in a single circular loop, for the special case in which the sphere is placed upon the axis of the loop. The arrangement is that shown in Fig. 1, but with the upper loop eliminated.

The force (F in dynes) with which the sphere is repelled from the loop is expressed in terms of the vertical axial distance (Z in cm) between the centers of the sphere and loop, amplitude of alternating current (I in rms amperes) in the loop, the radii (R_2 and R_1 in cm) of the sphere and loop, the frequency (ν in cycles/sec) of the alternating current in the *loop* and the resistivity (ρ in microhm-cm) of the sphere by

$$F=(3/50)\pi^2I_1^2G(x)A(y)(R_2/R_1)^3 \quad (5)$$

[4] See reference 2 (1939 edition), section 11.05. On p. 397 lines 11 and 12 read BD for D. In the 1950 edition, section 11.06 on the third line from the bottom of p. 398 read $p\to 0$ and on the bottom line both radicands should read $2/\pi x$. On p. 399 line 15 from the bottom read 12 for 13.

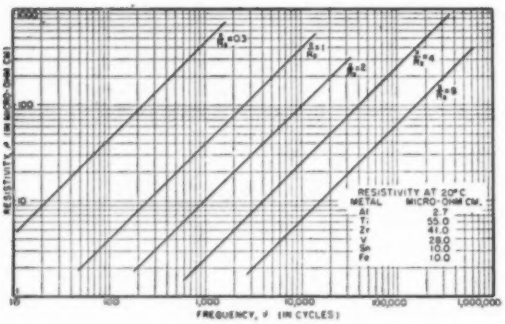


FIG. 4. Representation of parameter x for the evaluation of the dimensionless function $G(x)$ from Fig. 3 for use in formulas (5) and (6).

where

$$x(\text{dimensionless})=2\pi R_2((\nu/\rho 1000))^{\frac{1}{2}},$$

$$G(x)=1-\frac{3}{4x}\frac{\sinh 2x-\sin 2x}{\sinh^2 x+\sin^2 x},$$

$$y(\text{dimensionless})=2/R_1,$$

and

$$A(y)=y/(1+y^2)^4.$$

The quantitative behavior of the dimensionless function $G(x)$ is illustrated by Fig. 3 and values of x are conveniently determined from ρ, ν, and R_2 by Fig. 4.

Forces acting upon a 1-inch bronze ball (resistivity about 6 microhm-cm) at different points along the axis of a circular loop 4.85 inches in diameter, of $\frac{1}{4}$-inch copper tubing, were measured while a current of 600 amperes rms at 9600 cycles flowed briefly. There was good agreement with formula (5) as indicated by Fig. 5. In particular, a maximum repulsion of 0.50 percent of the weight of the ball was experienced at a distance of 1 inch from the plane of the loop. The same experiment with a 1.5 inch instead of a 1-inch bronze ball gave somewhat less close agreement. The measured maximum repulsion was $\frac{5}{6}$ of the calculated value in the latter case.

Formula (5) was extended by L. S. Castleman to the case shown in Fig. 1 of equal coaxial *loops* carrying equal alternating currents. The force in this case is

$$F=(3/50)\pi^2 I_1^2 G(x)B(y)(R_2/R_1)^3, \qquad (6)$$

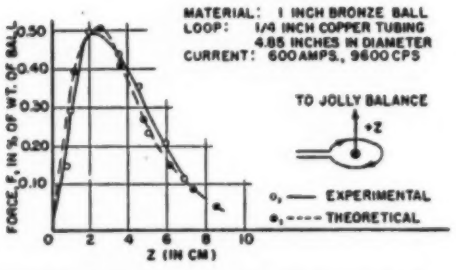


FIG. 5. Lifting force on solid metal sphere exerted by alternating magnetic field of a single circular loop along its axis.

where

$$B(y)=\left\{\frac{1}{(1+y^2)^{3/2}}\pm\frac{1}{(1+y_1^2)^{3/2}}\right\}\left\{\frac{y}{(1+y^2)^{5/2}}\pm\frac{y_1}{(1+y_1^2)^{5/2}}\right\}$$

and $y-y_1=S/R_1$, S (in cm) being the distance between the planes of the loops as illustrated by Fig. 1; Z now represents the distance above the plane of the lower loop, and Z and F are positive upward. The positive signs apply in $B(y)$ if the alternating currents have the same directions in both loops, the negative signs if the currents are opposite as in Fig. 1. Experiments were made as before with the 1-inch bronze ball and with coils as in Fig. 1, and there was the same order of agreement of measured and computed force. The greatest difference amounted to about 10 percent. The comparative results are illustrated by Fig. 6 when the currents in the loops are opposed and by Fig. 7 when they are not.

It is to be expected that such formulas as (5) and (6) become more precise as the sphere is made smaller, other things remaining the same. The above experiments are of value in indicating for how large a sphere the formulas retain satisfactory precision.

As regards stability with respect to vertical displacement, Fig. 6, in which current directions are opposed, depicts a stable situation and Fig. 7, in which current directions are alike, an unstable one. In Fig. 6 if the sphere is placed at the axis midway between the loops and is displaced slightly along the axis either up or down, a force acts tending to restore it to center. In contrast, in Fig. 7 the force would tend to move the sphere farther from center.

The situation in Fig. 6 is stable also with respect to horizontal displacement. This is shown by Fig. 8, which presents experimental results when the ball was displaced from the axis along a perpendicular to the axis at its midpoint. A formula similar to (6) could have been derived for this case, and no doubt the fit with

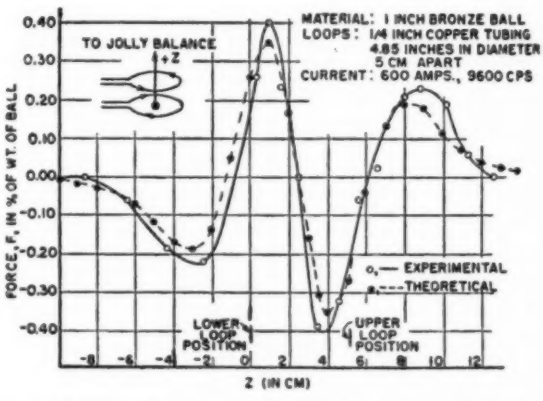


FIG. 6. Lifting force on a solid metal sphere exerted by alternating magnetic field of two coaxial circular loops (with current directions opposed as shown) along their common axis.

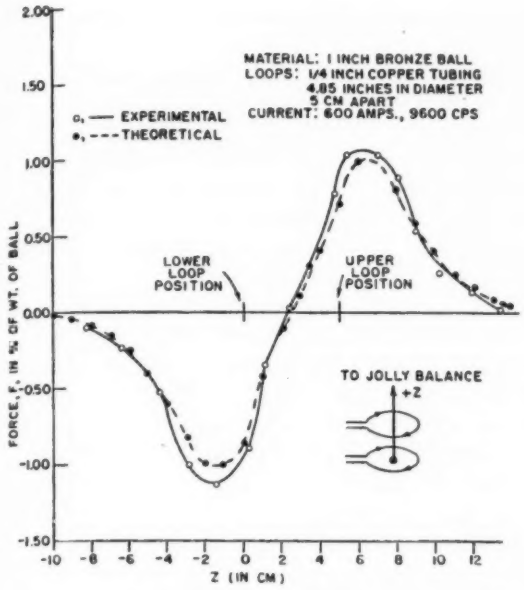


FIG. 7. Lifting force on a solid metal sphere exerted by alternating magnetic field of two coaxial circular loops (with current directions the same as shown) along their common axis.

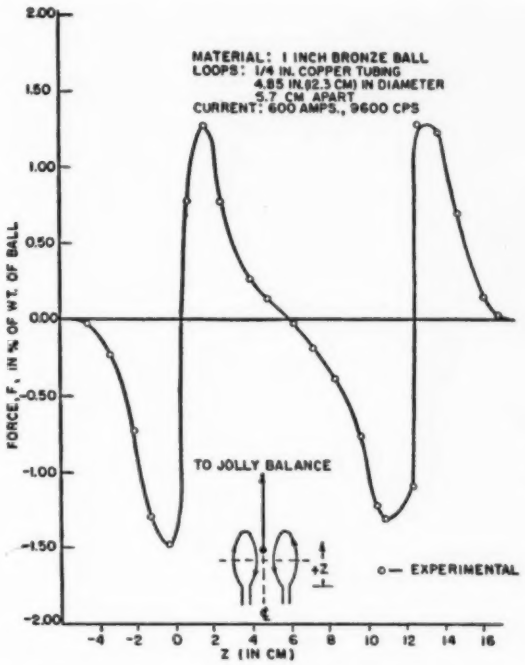

MATERIAL: 1 INCH BRONZE BALL
LOOPS: 1/4 IN. COPPER TUBING
4.85 IN.(12.3 CM) IN DIAMETER
5.7 CM APART
CURRENT: 600 AMPS., 9600 CPS
TO JOLLY BALANCE
+Z
o— EXPERIMENTAL
FORCE, F, IN % OF WT. OF BALL
Z (IN CM)

FIG. 8. Lifting force on a solid metal sphere exerted by alternating magnetic field of two coaxial circular loops (with current directions opposed as shown) along a loop radius projected onto plane midway between planes of the loops.

experiment would have been about as close in Fig. 8 as in 6 and 7.[5]

More generally, it would be possible in the same manner to calculate approximately the force exerted upon a sphere by currents in a set of coaxial loops which need not be equal in diameter. Any coil or pair of coils such as that in Fig. 9 might be approximated by such a set of loops. The magnitude of the force at any point, not necessarily on the axis, is proportional to the product of magnetic field strength B and the derivative of B in the direction of B, at the point. The direction of the force vector is that of the derivative. Computation might be simplified by assembling a set of tables of the components of B, and the components of the axial and radial derivatives of B, for the case of a single circular loop.[6] Coarse tables of this kind have been prepared.

Formulas for the force exerted by a *coil* upon a sphere would be of the same form as (5) and (6), that is, I^2 times $G(x)$ times a factor depending only on ratios of lengths. It follows that if a certain specimen can be levitated by a certain coil, then a specimen of twice the diameter, or eight times the weight, can be levitated by a coil scaled up by the factor 2 in all its dimensions, provided current is increased by the factor $2\sqrt{2}$.

If the coil is scaled up by a factor of 2 in all dimensions except that the size of the tubing remains unchanged, then twice as many turns of tubing can be wound into the doubled coil. In that case the current need be increased only by the factor $\sqrt{2}$. It may even be possible to incorporate not twice, but four times as many turns, since the area through which the turns are threaded becomes twice as large when the scale is doubled. Then the current would actually lessen by the factor $\sqrt{2}$, except for interference effects in a multi-layer coil. As for absolute values, it may be mentioned that 800 amperes at 9600 cycles has supported a half-pound bronze ball in a coil of largest diameter 9 inches.

The power dissipated[7] in the sphere varies with frequency in the same manner as lift varies with frequency, namely the manner specified in Eq. (5).

None of the above calculations touch the question of the stability of levitation in the molten state.

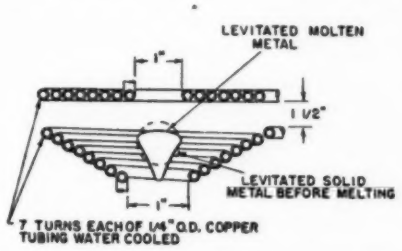


FIG. 9. Drawing illustrating levitation of solid and molten metal by a coaxial coil system, connected series opposition as in Fig. 1. The metal regardless of initial shape assumes a "top" form after melting.

[5] The forces in Fig. 8 are considerably larger than in Fig. 6. This suggests an arrangement in which two short solenoids face one another with a common horizontal axis and with current directions opposed, and the specimen floats between them. This was tried with Al. Stability of the solid was excellent. The melted Al took the shape of a wedge, with the edge down and at right angles to the axis of the solenoids.

[6] See reference 2, 1939 edition, section 7.04. In 1950 edition, section 7.10.

[7] Reference 2 (1950 edition), section 11.07 or 1939 edition, section 11.051.

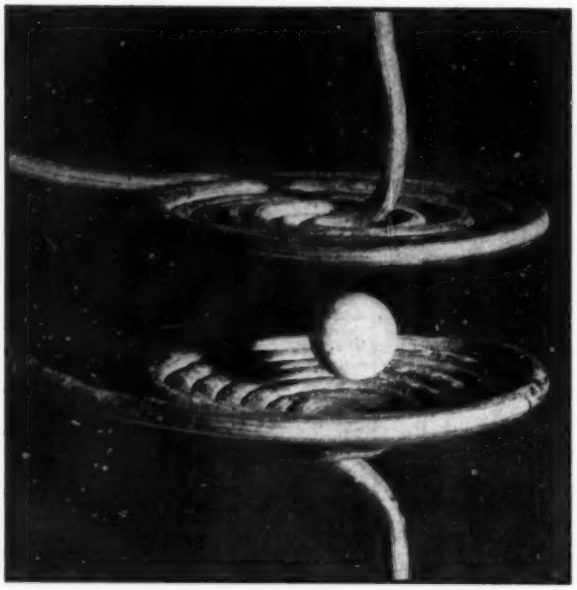


FIG. 10. Electromagnetic levitation of the solid Al sphere by the coaxial coil system of Fig. 9.

EXPERIMENTAL CONSIDERATIONS

The initial investigation at Bloomfield proceeded to the point where it was possible to insert from above or below the coil system slugs of any shape of various conductive materials into the space between coils ("interaction space") as illustrated by Fig. 9 for a particular coil system. Conductive materials such as brass, Al, Sn, compacted Ti powder, Mo, and others were used in the experiments. The solid specimen was levitated as illustrated by adjustment of the alternating current in the coils. The levitated specimen began to heat, and the rate of its temperature rise could be controlled by varying the magnitude of the alternating current, controlled at any rate above the level corresponding to the minimum current that would support the specimen. In those cases in which the specimen continued to be supported while molten, it assumed the shape of a top and remained stably levitated, but at a somewhat lower position in the field as illustrated. Sometimes Al already in the molten state was introduced into the interaction space; it immediately assumed the form illustrated by Fig. 9. In the levitated state the molten mass was vigorously agitated by the interacting alternating magnetic fields. The specimen tended to rotate about the axis of symmetry, but this rotation was generally slow and could be prevented by introducing axial asymmetry. Finally, the molten levitated mass could then be drained into a suitable receptacle, or dropped as a whole by merely adjusting the current in the coils. It may even prove possible to resolidify the levitated molten metal and then drop the solid into a suitable receptacle.

In an effort to obtain preliminary design data so as to provide more effective and stable levitation of the molten metal, many types of coil systems[1] were investigated.

Preliminary electrodynamic tests were also made with various coil configurations for evaluating their effectiveness in heating and levitation of metals. The levitated specimens were loaded by weights, and thermocouples were inserted to register average temperatures. Although this attack was susceptible to some defects, useful design data were obtained with the aid of electrolytic tank plots of the fields of a variety of coil designs. The data included the effects of leakage flux, coil configuration, electromagnetic coupling, magnitude and sense of the circulating current in the coils, and power input.

When molten metals were levitated in air, e.g., Al, it was evident that they soon became coated with oxides or other compounds. How important was this oxide coat in making molten metal levitation possible? On the one hand it seemed it could not be of much account for the vigorous agitation of the molten metal cracked whatever film was formed so that islands of film swam on the surface. On the other hand, stability of a melt with coils of the type in Fig. 9 might depend on preventing pour from starting at the bottom. A film of oxide there might possess an efficacy out of proportion to the area it covered. Since the chief applications of the method would of course be in a vacuum or a protective atmosphere where little or no film would form, the East Pittsburgh vacuum and atmosphere experiments to be described below were of major interest.

It is inherent in the coil designs[1] of Figs. 9–11 that the region of minimum levitation force exists at the coil axis. This results in the tendency to drip. Two schemes may be employed to alleviate this difficulty. The first would be to create the field with current sheets or a section of torus solenoid. The second solution is to move the minimum at such a rate and orbit that the inertia of the molten metal prevents the tendency to drip. The latter may be accomplished by mechanically or electrically moving the induction field. Single and polyphase systems appear feasible. Experimentally, only the single phase system was tested. The molten levitated metal assumed an approximately spheroidal form. Even though improved results are produced by this scheme the complication of the coil system is a serious disadvantage. Electrical rotation of the alternating magnetic field for levitation of the molten specimen is preferred over mechanical, but the latter may have some advantage in the simplicity of the coil system.

Figure 10 illustrates the coil system of Fig. 9 and shows the levitation of an Al sphere in the interaction space by the induction fields of the two coils connected in series opposition. Appreciable restoring forces were experienced when the sphere was pushed in various directions from its equilibrium position in the interaction space. In each case the sphere returned to its

equilibrium position near the center of the coil system as illustrated by Fig. 10.

The Al specimen was allowed to start melting in air while levitated by the alternating magnetic field of the coils. As it did so it assumed the shape of a top as illustrated by Fig. 11. Sometimes, as melting of the levitated specimen progressed, portions thereof flowed downward and intermittently dropped out of the field. In order to avoid this tendency to drain, the size and mass of the levitated metal may be decreased, or the circulating current in the coils may be increased, or the coils may be more effectively shaped in order to levitate both the molten and solid metal. All of these expedients were tried and found to stop intermittent drainage.

Experiment indicates that for most effective levitation and induction heating of a metal specimen an optimum coil configuration exists and that it varies with the type of specimen and whether it is in molten or solid state. Although data have been obtained regarding the levitation characteristics of various coil configurations, it is only a preliminary stage and not sufficiently established for general application. However, it is appropriate to cite a typical experiment.

The example involves a double coil system illustrated by Fig. 9, the coils connected in series opposition as illustrated by Fig. 1. Stable levitation was observed with solid and molten Al (21.5 grams), solid Mo (154 grams), and solid Cu (80 grams). Both of the latter elements were at red heat. The coil currents amounted to 630, 720, and 630 amperes, respectively. No attempt was made to melt the Mo and Cu specimens. A $\frac{1}{8}$-inch diameter, 1-inch long Al rod was fed into the molten metal from above the coil system of Fig. 9 without disturbing the stability of the levitated molten Al. This indicates feasability of replenishing the melt.

Some attention was paid to levitation of solid spheres in solenoids. Such coil configurations may prove on further investigation to be as satisfactory as, if not superior to, the other arrangements cited in this paper.

Coils employed at East Pittsburgh were similar to some of those at Bloomfield. Currents ranged up to about 1200 amperes. Frequency was always 9600 cycles at East Pittsburgh, 10,000 cycles at Bloomfield.

In air, in the solid state, various masses of Al, Ag, and aluminum bronze were levitated stably just as at Bloomfield. In the molten state, Al was levitated with ease. Molten Ag was levitated between coaxial coils on one occasion, but drifted laterally out of the field. On other occasions, when melting commenced, the molten silver drained downward and dripped away from the lower surface of the remaining solid metal, apparently without much interference from the alternating magnetic field which was concurrently supporting the solid metal. It is to be observed that Ag melts in air without forming a protective oxide film like Al.

In a vacuum of 10^{-4} or 10^{-5} mm of Hg, there was no difficulty in levitating Al, Ag, and Ti in the solid state.

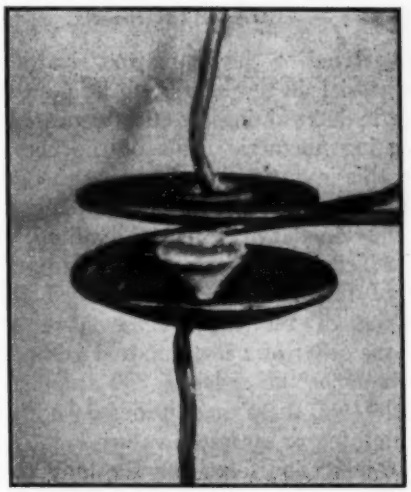


FIG. 11. Electromagnetic levitation of molten Al by coil system of Fig. 9. Note the "top" shape of the suspended molten metal.

On melting, however, Al usually dripped downward, even in a coil and with a current that would have supported the liquid if surrounded by air instead of vacuum. In the end it was found possible to levitate approximately 10 grams of molten Al in vacuum for about one minute. But even then it became apparent that there was still a film which segregated as wrinkles at preferred sites on the molten surface in about one minute. This was a trial in which an Al ball in vacuum, as in air, was levitated and heated in the solid state and levitated in the liquid state as already described. For example, an Al ball 1 inch diameter was floated by means of currents of 600 to 800 amperes. It split in half, but one half continued to float, melted, then floated molten for about a minute in the shape of a thin pear with point down. The surface appeared very bright and clean at first, but toward the end traces of film were seen. The gauge indicated a vacuum of 10^{-5} mm Hg throughout the run.

A lump of Ti was almost completely melted while in levitation. Ag, however, always began to drain as soon as it melted. At the same time opaque films began to form on the inside surface of the bell jar which enclosed the vacuum, so that observation of melted silver could never be carried on long.

In another approach to the problem, a pair of parallel conductors were employed instead of coils. The conductors were 6-inch lengths of $\frac{3}{8}$-inch o.d. copper tubing, with a 1-inch space between. Currents, in phase conjunction, reached about 2000 amperes. Bars of Al, Zr, and Cu, about $\frac{1}{2}$ inch in diameter and 4 inches long, were readily levitated in the solid state between the conductors. In hydrogen atmosphere (to eliminate oxide coating) a Cu bar was levitated partially molten. However, support of completely molten metal was not achieved by this arrangement.

CONCLUSIONS

Factors related to the shape and size of solid conductive material do not appear to be paramount considerations in determining the effectiveness of levitation by alternating magnetic fields of the character considered in this investigation. The mass to be levitated governs, at a given frequency and coupling, the magnitude of the circulating current required, the availability of which is in turn dependent in part upon the efficiency of the coil system and tuned circuit as a whole. Although the experiments have dealt with masses of the order of a few hundred grams and circulating currents of the order of 700 amperes at about 10,000 cycles/sec, these magnitudes do not necessarily represent limiting or optimum values.

Both the levitation force and the induction heating power input vary as the square of circulating current in the coil system. The resistance of the coil system should be a minimum, and the frequency and the equivalent Q of the levitated mass should be large enough so that the quantity

$$(Q^2/(Q^2+1))(M/L)^2 \rightarrow (M/L)^2,$$

consistent with a tolerable rf radiation loss. M denotes the mutual inductance between the coil system and specimen. L denotes the equivalent inductance of the specimen. On a scaling basis it has been estimated that masses of several pounds may be stably levitated with a 50 kva, 10,000 cycle source.

If electromagnetic levitation can be expanded to a larger scale, the following advantages can be anticipated:

1. The specimen touches no crucible or container during the heating, melting, and draining stages.
2. The heated or molten specimen can be protected by a suitable atmosphere or a vacuum.
3. Volatile impurities can be distilled or pumped away.
4. The molten specimen can be drained gradually or dropped as a whole, and it may even be feasible to solidify the still levitated melt and then drop it solid.
5. The melt is thoroughly mixed electromagnetically.
6. The specimen can consist of metal powders and alloying constituents mixed and sintered together or bound by a suitable binder.
7. Alloying additions can be made while the melt is in levitation.

The limitation of the existing coil systems is the relatively low electrical heating efficiency of induction coils surrounding a relatively small specimen.

It cannot be said that the feasibility of levitating molten metal in a protective atmosphere or in vacuum has been firmly established. The nearest approaches were the experience in a coil in vacuum with the Al ball that split, and between parallel conductors in hydrogen with the copper bar that partly melted. It seems certain that oxide coats which form in air make levitation easier, but whether such coats are essential, at least for specimens beyond a certain size, remains to be investigated further. The successful experiments with Al using rotating alternating current fields may offer the solution to levitation of other metals in vacuum.

ACKNOWLEDGMENT

We wish to take this occasion to acknowledge our obligations to those who aided us in realizing a measure of our object. Our appreciation is due to E. G. F. Arnott, Director of Research at Bloomfield, for encouragement and assistance, to E. Gregory and J. G. Jacobsen for contributions to the experimental work, and to L. F. Ehrke and W. Adamus for photography. We are most grateful for the contributions at East Pittsburgh of J. A. Hutcheson, Vice President, C. M. Slack, the late R. C. Mason, K. McCoy, J. W. Salatka, and C. R. Simcoe. Special acknowledgment is due to L. S. Castleman, who derived Eq. (6) and made comparisons summarized in Figs. 6–8.

Arcing at Electrical Contacts on Closure. Part III. Development of an Arc*

L. H. GERMER AND J. L. SMITH
Bell Telephone Laboratories, Incorporated, Murray Hill, New Jersey
(Received February 19, 1952)

A description is given of a system made up of experimental electrodes and an oscilloscope by means of which the potential across the electrodes can be recorded with a time resolution of about 10^{-9} sec and a potential sensitivity of 1-trace width per volt. The closure of the electrodes to produce a short arc is synchronized with the oscilloscope sweep so that the beginning of the arc is photographed.

As an arc starts the potential across the electrodes decreases more or less gradually from the applied voltage to a steady value characteristic of the metal of the electrodes. The course of this change is extremely variable as is also the time over which the change is spread. The average value of the time appears to vary with circuit inductance and with the nature of the electrode surfaces. For inactive silver surfaces and an inductance of 0.10 μh the average value of the time is about 0.007 μsec, and for active surfaces and the same inductance 0.011 μsec. For active surfaces and an inductance of 5 μh the average value of the time is 0.02 μsec.

The electrode separation at which an arc strikes is determined from the oscilloscope traces and from a correction for the height of the mound of metal thrown up by the arc. For active silver electrodes the average separation (at 40 or 45 volts) corresponds to a gross electric field of 0.8×10^6 volts/cm, and for inactive silver electrodes to a field of 2.3×10^6 volts/cm. These are probably better values than earlier measurements of these fields. There has not yet been any success in interpreting these phenomena in terms of fundamental processes.

IF the length of an electric arc is sufficiently small and its lifetime not too short, the potential across the arc for most of its life is determined by the electrode metal. It is independent of current, the separation of the electrodes, and whether or not there is gas between them.[1] When current starts to flow at the very beginning of one of these "short arcs" the potential across it is often above the steady characteristic value and approaches this value only after an appreciable time. This conclusion regarding the gradual approach to the arc potential characteristic of the electrodes was reached indirectly without observation of the course of the potential during this formative period.[2] The gradual change has not been directly observed previously. The present paper is an account of direct measurements of the potential during this period. The results confirm the indirect earlier observations; the formative period is found to very greatly in length with a maximum value of about 0.05 microsecond, quite well in line with the earlier conclusion that it is of the order of 0.02 microsecond.

Since these short arcs avoid some of the complexities of longer arcs it is to be hoped that study of them will be rewarding in providing data which can be used to gain a theoretical understanding of the development of an arc. Up to the present time there has been no success in interpreting in terms of fundamental processes the highly variable way in which the arc potential changes before its steady value is established. The experiments are presented here in the hope that interpretation will come later.

One of these short arcs is produced by bringing two electrodes toward each other to discharge a small capacity. Whether or not an arc strikes before the electrodes touch depends primarily upon the nature of the electrode surfaces and upon the inductance and resistance in the circuit.[1] All of the experiments reported here have been made upon silver electrodes discharging a condenser charged to a potential difference of 35, 40, or 45 volts. Two cases are to be sharply separated; in one the metal surfaces are moderately clean, "inactive" as we have described them previously; in the other they are partially covered by carbonaceous material, "active" in our nomenclature.[1]

The experiments lead to information of two sorts. Primarily, they accomplish what they were intended to do—discover the way in which the potential between two electrodes decreases during the time the arc is being established, from the initial potential to the final constant value characteristic of the metal of the electrodes.

The experiments give also new sets of measurements of the electrode separations at which short arcs strike. Interest in such measurements arises partly from the fact that a prevalent theory for the initiation of an arc at contact closure in a low voltage circuit assumes that the first step in the development of an arc is touching at a small point; if this first point of contact is burned out by the rush of current through it, an arc may be established in the resulting vaporized metal. It has been believed that there can be no closure arc in a low voltage circuit if the first point of contact is so substantial that it is not burned out.[3] The measurements reported here confirm previous observations[4,5] that this is not at all the usual way for a closure arc to start, at least at the currents and voltages of these experiments. In most cases the arc is initiated by field emission current when the electrodes are still relatively far apart. Often the arc is over before they touch for the first time. Because

* Presented before the American Physical Society, November 3, 1951.

[1] L. H. Germer, J. Appl. Phys. **22**, 955 (1951), Table II.

[2] L. H. Germer, J. Appl. Phys. **22**, 1133, 1138 (1951).

[3] R. Holm, *Electric Contacts* (Hugo Gehers, Stockholm, 1946), p. 280.

[4] L. H. Germer and F. E. Haworth, J. Appl. Phys. **20**, 1085 (1949), pp. 1090–1092.

[5] See reference 2, page 1136.

of the high resolving power of the oscilloscope used in the latest tests, the new measurements establish this conclusion with a degree of certainty beyond that of the earlier work.

EXPERIMENTAL APPARATUS

Oscilloscopic records of the potential across an arc between two closing electrodes have previously shown an initial perfectly abrupt decrease in the voltage. The gradual change from the applied potential to that characteristic of the electrodes, which has been postulated to account for some earlier observations,[2] has not been previously observed because of inadequate resolution of the oscilloscope. For example, the oscilloscope trace of the left-hand side of Fig. 1A of reference 1 shows an arc lasting for 3 μsec at a constant potential of 11 volts without any indication whatsoever of the way in which the voltage changes from the applied value of 35 to the steady value of 11. That the change should appear to be perfectly abrupt on this trace is understandable from the fact that about 0.3 μsec is required to start the oscilloscope sweep, whereas the predicted duration of the transition is of the order of 0.02 μsec. An oscilloscope is required having much better time resolution.

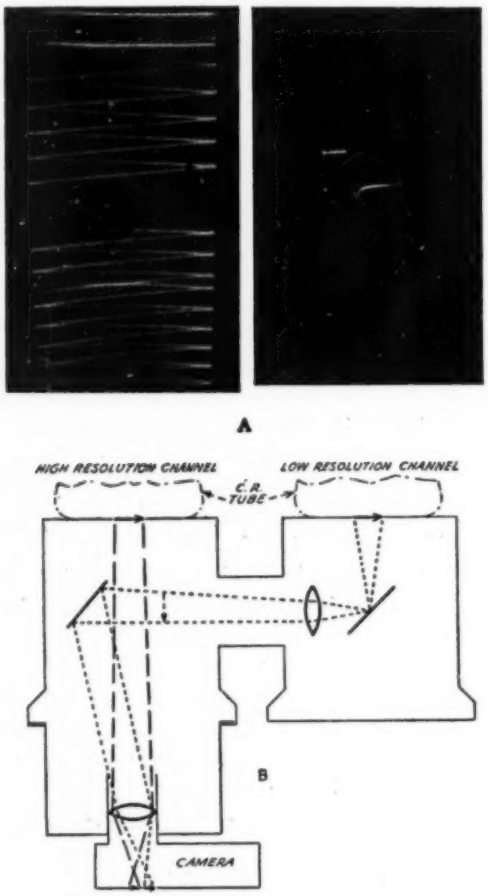


FIG. 1. Schematic diagram of optical system for photographing a low speed and a high speed trace of the same event, with a representative pair of traces.

A satisfactory oscilloscope has been built around a DuMont 5XP11-A tube[6] having a metallized screen. The leads to the plates of this tube come out of the narrow neck of the tube and go directly to two silver contact wires one of which is soldered to the armature of a loudspeaker unit so that contact between the silver wires is broken when a potential is applied to the magnet of the unit and contact is re-established after the potential is removed.

Each of the two lead wires from one of the oscilloscope plates to one of the contact wires (about 1 mm from the contact itself) is about 8 cm long. The natural frequency of the circuit from one oscilloscope plate to one of the silver contact wires, across the contact to the other wire and back to the other oscilloscope plate has been measured to be 330×10^6 cycles per second which corresponds to an LC product for this circuit of 2.3×10^{-19} henry farad. The plate capacity is given by DuMont as 4 $\mu\mu$f. Our measurement of the capacity of the plates plus that of lead wires out to the point where a damping resistance is to be inserted in the circuit is 5.8 $\mu\mu$f. The capacity which is effective in determining the natural frequency of the circuit is probably closer to the former of these values. From this one calculates for the circuit inductance $2.3\times10^{-19}/4\times10^{-12}=0.057$ μh, and for the critical damping resistance 240 ohms. A resistance of 200 ohms was inserted in this oscilloscope plate circuit and has been used in all measurements reported here.

The electron beam is swept across the screen of the oscilloscope tube in about 0.1 μsec. Before deflection the beam is accelerated by 2 kv and the post-deflection acceleration potential is 25 kv. The latter high value is necessary in order to have sufficiently high light intensity for photography. This is adequate with a lens of aperture 1.5 and linograph panchromatic film used with Eastman D19 developer.

The oscilloscope sweep is initiated by the same potential which is applied to the speaker unit to bring the contacts together. On its way to the tripping mechanism the suddenly applied potential goes through an adjustable delay circuit so that the electron beam can be made to begin to move just before the start of an arc between the approaching contacts. The effective length of path of the beam on the screen is increased by a factor of about 30 by sweeping it back and forth across the screen and at the same time deflecting it vertically about 15 or 20 times more slowly, in what the television people call a raster sweep. The horizontal sweep is obtained by electrostatic deflection of the beam by a 5-megacycle triangular voltage wave of about 50-volt peak-to-peak amplitude applied across the second pair of deflecting plates, and the slower vertical sweep by

[6] F. W. Kammerer of the Bell Telephone Laboratories contributed to the design of this tube, as well as to the design of the oscilloscope in which we have used it.

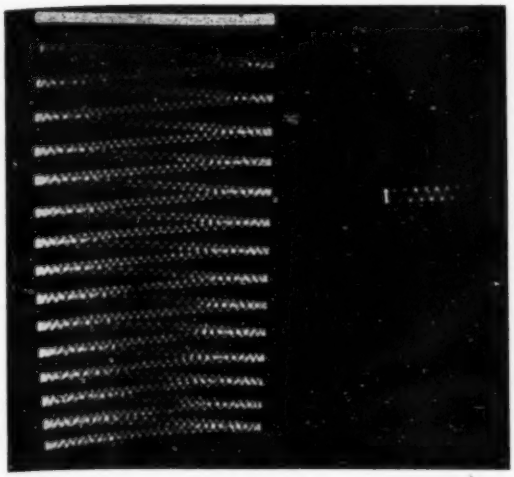


FIG. 2. Calibration by two standard frequencies, 295 Mc/sec and 2.00 Mc/sec.

magnetic deflection by current through a pair of coils mounted on the narrow neck of the tube. This current comes from a 0.01-μf condenser charged to 400 volts which is shorted through the coils by an R.C.A. 3D22 thyratron triggered by the potential which starts the electrostatic deflection.

There are four reasons for increasing the length of sweep as just described. These are: 1. Lack of reproducibility of the time interval from the activation of the electromagnetic unit to the beginning of the arc. 2. The fact that in many cases the arc itself, or repeated arcs, last for more than a single sweep time of 0.1 μsec. 3. The time from the initiation of the sweep until the unblanking circuit brings the beam intensity up to full value is greater than 0.1 μsec. 4. The need of seeing events long subsequent to the arc under study, for example the final closure of the electrodes, without which the phenomena under investigation might not be interpretable.

Even although the total length of the recorded trace on the oscilloscope screen is increased to about 30 transits, 3 μsec, it often and perhaps usually turns out that a still more extensive and a more compact record of a particular closure is essential. Without an oscilloscopic record at low sweep speed many of the high speed traces would be interpreted incorrectly, and no meaning at all could be discovered for others. To satisfy this need the potential across the experimental contacts is traced simultaneously at lower sweep speed upon the screen of a second oscilloscope, and by means of two mirrors and a lens the low resolution trace is "relayed" so that it too enters the camera lens and appears on the photographic record side by side with the high speed raster trace. A schematic diagram of the optical system and a pair of typical traces appear as Fig. 1.

On Fig. 2 are recorded traces of two calibration frequencies applied respectively to the plates of the two tubes. These frequencies of 295 and 2.00 megacycles per second permit one to calibrate the time scales. It is clear that on the high speed trace the speed of the beam is appreciably different when traveling to the right and to the left, and varies a little along each trace. These variations can be taken into account in interpreting any particular pattern. For many purposes it is satisfactory to assume for the time of a single sweep across the screen the mean value of 0.096 μsec.

All of the tests reported here are records of the potential across a pair of silver contacts as they discharge a capacity of 391 $\mu\mu$f, charged to 35, 40, or 45 volts, through various inductances in the range from 0.096 μh to 1000 μh. The circuit consisting of the capacity C, the inductance L and the contacts (Fig. 3) is built with small physical dimensions and is isolated from the low resolution oscilloscope channel and from the battery circuit by high resistances (R_1 and R_2 of Fig. 3), and from the plates of the high resolution oscilloscope by the damping resistance of 200 ohms described above. The electromagnet opening and closing the contacts is actuated by a multivibrator about once in two seconds.

Photographs are taken on a roll of film in a Leica camera with an auxiliary iris diaphragm shutter of large aperture. The shutter is opened by hand at an indicating click in the multivibrator system, and it closes automatically half a second later.

In order that the contacts can be activated when desired by operating them in an atmosphere of an appropriate organic vapor,[1] they are enclosed in a small plexiglas case which is attached to a large brass chamber having a total volume of 33.2 liters. This large volume is desirable so that a particular desired low vapor pressure can be established around the contacts by using an amount of the liquid compound which is not too small to measure accurately.

CLOSURES OF INACTIVE AND ACTIVE SILVER ELECTRODES

Active electrodes are distinguished from inactive electrodes by the maximum circuit inductance through which an arc can strike when they are brought together.[1] Since the limiting inductance for inactive electrodes is of the order of 1 μh, the distinction has been made for

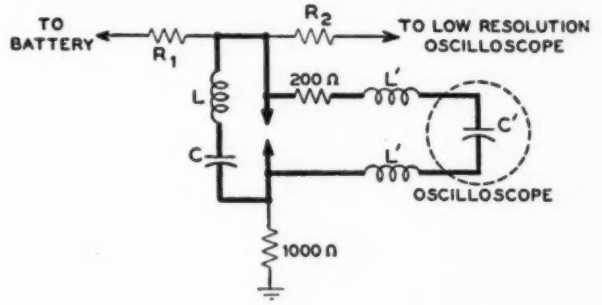


FIG. 3. The experimental circuit. The contacts are represented by arrows. $C=400$ $\mu\mu$f for all tests. $C'=4$ $\mu\mu$f, but in an oscillation with the contacts separated C' is replaced by an effective capacity $C''=12$ $\mu\mu$f. The total inductance $L'=0.06$ μh. In various tests L has the values 0.10, 5.0, and 48 μh. $R_1=10^6$ ohms. $R_2=10^4$ ohms.

the purpose of the present experiments by tests in a circuit containing an inductance considerably higher than this, 5.0 μh. Activation of the silver electrodes has been accomplished by opening and closing them repeatedly in an atmosphere containing an appropriate organic vapor in a circuit in which they discharge a small condenser on each closure but break only a current which is quite negligible. The activating compound is *d*-limonene and the partial pressures which have been used in the present experiments have been in the range between 0.054 and 0.010 mm of mercury.[7] In experiments with active electrodes and circuit inductance other than 5.0 μh it is desirable to test at 5.0 μh after every test at the inductance being studied. For this purpose the chamber containing the test contacts contains an external switch by which the circuit inductance can be changed from 5.0 μh to any other value under investigation.

The contacts are brought together to discharge a small condenser through various small inductances. The same condenser is used throughout. It has a measured capacity of 391 $\mu\mu$f, but we have made a small and reasonable allowance for unavoidable distributed capacity by assuming the capacity to be 400 $\mu\mu$f. The largest inductance for any of the traces which are reproduced here is 48 μh which was directly measured. Lower values of inductance have been determined by observing the natural resonant frequency of the circuit with the contacts closed and calculating the inductance from this frequency. For contacts in the inactive condition the condenser is attached to the contacts through short wires with no added inductance or through a very small inductance. The natural frequencies under these conditions are 26 mc and 16.5 mc from which the inductances are determined to be 0.096 μh and 0.23 μh, respectively. With contacts in the active condition the minimum wiring between condenser and contacts is necessarily slightly longer because of the switch by which the inductance of 5.0 μh is inserted into the circuit to test for activation, and natural frequency of the circuit when it has no added inductance is 25.2 mc corresponding to an inductance of 0.100 μh. The inductance of 5.0 μh is itself determined also by observing the natural circuit frequency and calculating the inductance from the known capacity of 400 $\mu\mu$f.

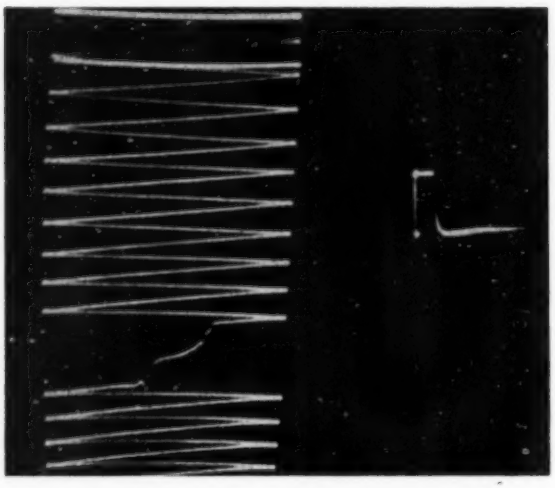


FIG. 4. Active silver electrodes discharging through an inductance of 0.10 μh a 400 $\mu\mu$f condenser charged to 45 volts. The arc between the electrodes occurs when they are still far apart and the high resolution trace runs off the screen before the contacts touch, with a reverse potential difference of about 5 volts between them.

The potential across the contacts as they are brought together is recorded in the traces of Figs. 4, 5, 8, and 9. The complete photographic record is shown for the closure of Fig. 4 only. For each of the other closures the photographic print has been cropped to leave only the essential part; this is for the sake of compactness in the reproduction. Although the low resolution trace seems not to help in understanding these closures, without it we could not have been confident of the interpretation. In many cases one finds from the low resolution trace that the apparently obvious interpretation of the high speed trace is not right, and in other cases one is unable even to think up a plausible interpretation without the low speed trace, when with this trace it is perfectly clear.

The calculated duration of an arc for the inductance of 0.10 μh is $\pi(LC)^{\frac{1}{2}}=0.020$ μsec which agrees excellently with the observed arc time on the pattern of Fig. 4 for active electrodes and of Fig. 5(A) for inactive electrodes. For each of these closures the arc potential decreases during the entire arc time just about reaching the steady value characteristic of silver electrodes at the time the arc goes out. This agreement between the arc time and that taken for the characteristic potential to be reached is merely a coincidence; for other closures of active silver electrodes with larger circuit inductances the steady arc potential is established in a time which may be a small fraction of the total arc life, as in earlier published patterns (e.g., Fig. 1B of reference 1) and Figs. 8 and 9.

There are two important differences between the patterns of Fig. 4 and 5(A), in addition to that caused by the slight and incidental difference between the potentials of the charged condenser. The first of these is the difference between the way in which the potential decreases at the very beginning of the arc. For the active electrodes the potential across the arc does not have the initial precipitous drop shown for the inactive electrodes. This sudden decrease of voltage is exhibited by all arcs between inactive electrodes. Between active

[7] Limonene is a major constituent of turpentine. In the large tightly closed 33.2 liter chamber containing the contacts the pressure of 0.054 mm corresponds to a total limonene content of 0.020 cc and 0.010 mm to 0.0037 cc. The latter pressure is about the lowest at which silver electrodes can be activated and is much below the lowest limonene pressure which is effective in activating palladium electrodes (reported as 0.06 mm Hg, L. H. Germer, Bell System Tech. J. 30, 933 (1951), p. 935). It has been stated earlier that the lowest limonene pressure for the activation of silver electrodes is not much below 0.025 mm (see reference 1, Table I).

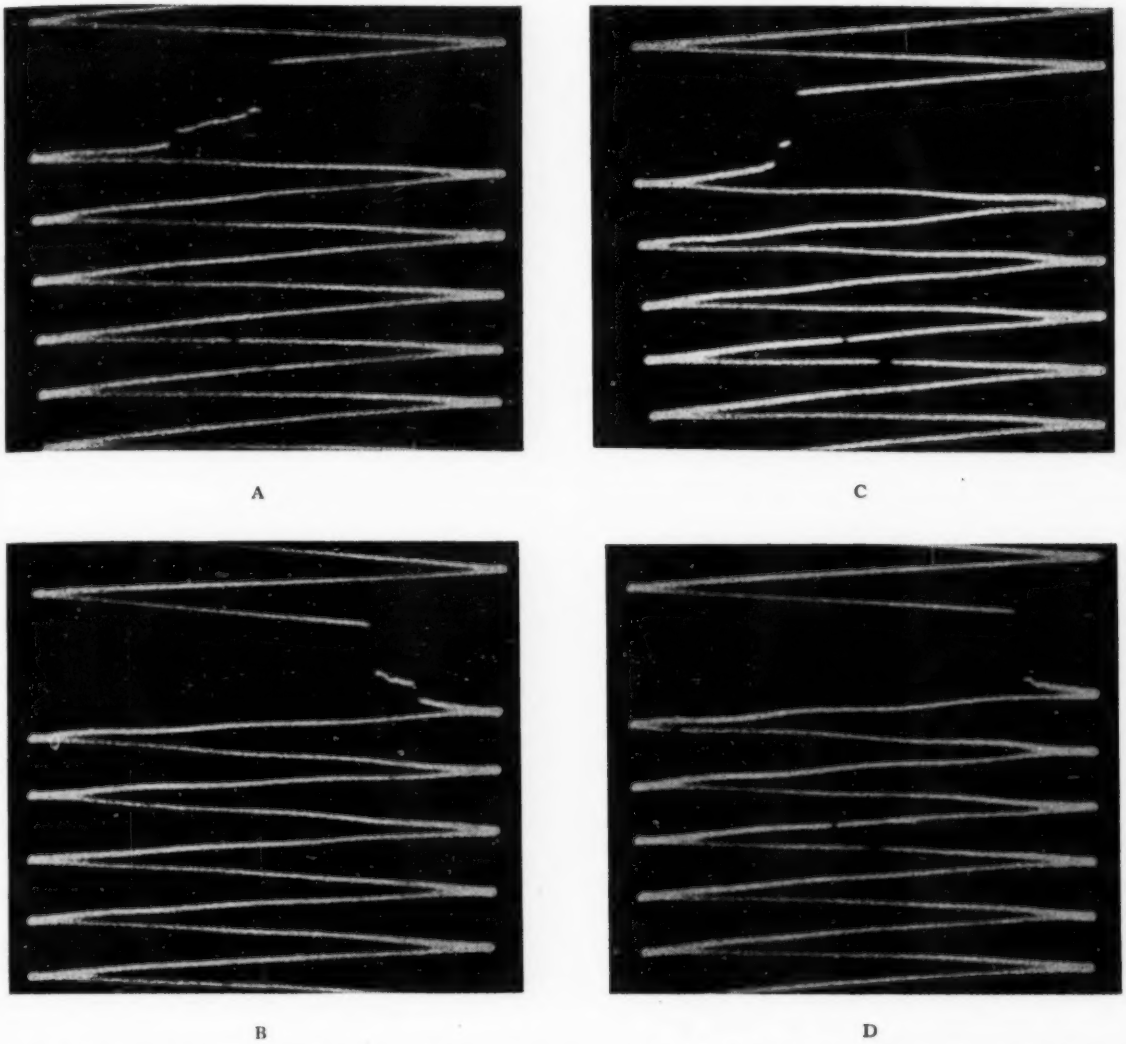


FIG. 5. Inactive silver electrodes discharging through an inductance of 0.10 μh a 400 $\mu\mu$f condenser charged to 35 volts. A. Closure of the contacts just at the end of the arc. B and C. Closure before the end of the arc, with the contacts shorted by metal thrown up by the arc. D. Closure in about 0.001 μsec, without any observable arc.

electrodes the potential often changes gradually even at the very beginning (as in Fig. 4), although on the other hand many traces of closures between active electrodes exhibit an initial drop like that of Fig. 5(A). Inactive electrodes always show the abrupt initial fall of potential, but the potential across active electrodes sometimes changes gradually at the beginning and sometimes abruptly.[8]

Critical Electric Field

A second important difference between the closures represented in the traces of Figs. 4 and 5 is in the separation of the electrodes at which the arc strikes. The arc of Fig. 4 between active electrodes occurs when they are relatively far apart and the high resolution trace runs off the screen before the electrodes touch, the potential difference between them at the end of the observable high speed trace being about 5 volts in the direction opposite to that of the initial voltage. The inactive electrodes, however, touch just as the end of the arc. In fact inactive electrodes often touch before an arc is over, as in the closures represented by the traces of Figs. 5(B), 5(C), and 5(D). This is in fact most usual for inactive electrodes with a potential of 35 volts, but very rare for active electrodes. It is clear that the electrode separation at which an arc strikes is much greater for active electrodes than for inactive. This is the conclusion reached in earlier experiments (see reference 2, Table I), but the present tests give better results and are more reliable.

Statistical data on the separation of electrodes at which an arc strikes have been obtained from measurements upon many oscilloscope traces. For this study

[8] The extreme sharpness of the potential change across the oscilloscope plates which is sometimes observed is proof that the calculated time resolution of the oscilloscope of about 10^{-9} sec is actually attained.

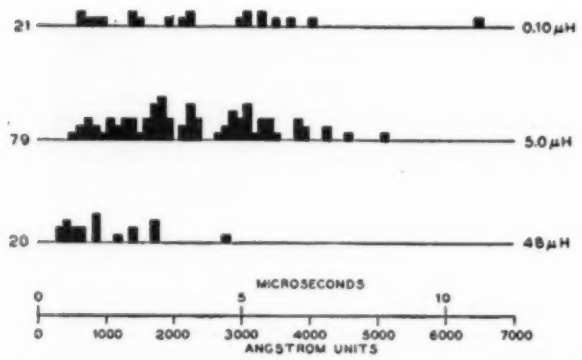


FIG. 6. The distribution of the time from an arc to closure for active silver electrodes, with also a scale of distances obtained from the time by multiplying by the electrode velocity 6.0 cm/sec.

63 rolls of film have been used with records of over 2000 closures. For all of these tests the closure velocity of the contacts has been adjusted[9] to be 6.0 cm/sec so that observed time from the arc to closure can be transformed into electrode separation. On some records the contact closure appears on the high speed trace, on other records on the low speed trace, but on a large number there is no clear closure record so that these traces are valueless for this study. Many traces are faulty in other ways, such as synchronization failure and double traces, so that less than 10 percent of the traces have been used in the study.

The distribution of elapsed time between an arc and contact closure for 120 closures of active silver electrodes is given in Fig. 6. Data have been separated into three different distributions, for circuit inductances of 0.10, 5.0, and 48 μh. At the left of each distribution is given the number of observations which are plotted, respectively 21, 79, and 20. An abscissa scale of separation in angstrom units has been added by multiplying the observed times by the electrode velocity of 6.0 cm/sec. For the two smaller values of inductance the over-all average separation seems to be about 2500A but the spread is very wide around this. The data for closure in the 48 μh circuit indicate a significantly smaller average separation, but this is probably attributable to discrimination in the selection of data so that these observations should not be used in obtaining an average value of electrode separation. The disturbing factor is the fact that in the 48 μh circuit a closure arc develops in most instances only after repeated discharges of the oscilloscope plate capacity, as in the trace of Fig. 9(A) discussed below. Only the less usual closures in which the arc develops without such preliminary discharges, as in Fig. 9(B), have been measured up for plotting in Fig. 6. It seems probable that at the first discharge an arc is more likely to develop at once if the separation happens to be small than if it is larger; if this is so, then the data for 48 μh exhibited in Fig. 6 have been selected so that they represent an unfair proportion of those closures at which the separation at the first arc was below the average value.

It has already been pointed out that when there is an arc between inactive electrodes they close in most cases before the arc has run to its full time. In fact, closure often occurs before the arc has lasted as long as 0.005 μsec. By multiplying by 6 cm/sec, this time of 0.005 μsec gives an electrode separation of $3A$ which is obviously in error. The fault almost certainly comes in from neglecting the mound of metal which is thrown up by the energy of the arc. Because of the appreciable height of this mound the electrodes close more quickly than they would if the surface had not been disturbed. If closure occurs at any time before the arc should naturally end, that is at a time t' which is less than $\pi(LC)^{\frac{1}{2}}$, it is to be presumed that the closure is produced by a mound of metal thrown up by the arc, and the separation of the electrodes at which the arc struck is simply the height of the mound. In earlier work this height has been directly observed and we can now make use of these earlier measurements to convert time into electrode separation.

The average height of a mound of metal thrown up by an arc upon platinum electrodes has been determined to be $h=2.2\times10^{-5}E^{\frac{1}{3}}$ cm when the energy E is measured in ergs.[10] The energy in an arc is $E=\int_0^{t'} vidt$. If we neglect the fact that the arc potential v is not constant this becomes $E=v(V_0-v)C(1-\cos\epsilon\pi)$, where V_0 is the initial potential across the condenser C and ϵ is the proper fraction obtained by dividing the time the arc lasts (until the contacts are closed by the metal mound) by the full arc time $\pi(LC)^{\frac{1}{2}}$. The separation of the electrodes at which the arc strikes is then

$$h=4.7\times10^{-3}[v(V_0-v)C(1-\cos\epsilon\pi)]^{\frac{1}{3}}\text{ cm}. \quad (1)$$

All of the available data for the closure of inactive electrodes have been plotted in Fig. 7. There are only five cases in which an arc between inactive electrodes does not result in closure at or before the normal end of the arc. These cases are plotted as black squares, the abscissa scales being measured time in microseconds and the corresponding calculated distance in angstrom units. There are 61 cases of closure at or before the normal end of the arc when the inductance is 0.10 μh, and 12 cases of this sort for inactive arcs in a circuit with $L=0.23$ μh. These 61 cases and 12 cases are plotted as shaded areas at the left of the figure; there is a linear abscissa scale of calculated separations in angstrom units with the total energy of a full arc $2v(V_0-v)C$ converted into a separation of 2800A by Eq. (1) and plotted to coincide with the zero of the right-hand scale. At the left are given also two nonlinear scales on which the shaded areas give the distribution of the arc times before closure in fractions of a microsecond.

The data leading to the numerical constant of Eq. (1)

[9] As described on page 1133 of reference 2.

[10] See reference 2, footnote on page 1136.

were obtained for platinum electrodes and are being improperly applied here to silver. Another inconsistency is the fact that the arc voltage has been assumed to be constant in the derivation, whereas the main burden of the present paper is the investigation of the way in which the arc voltage deviates from constancy. We believe, however, that these faults are relatively unimportant numerically because of the cube root in Eq. (1).

In looking at the distributions of Fig. 7 one is surprised by the peak in the lower curve just at 0.020 μsec, at 2800A according to the way time has been converted into distance by means of Eq. (1). There seems to be no reason why an arc should tend to strike between approaching electrodes just when they reach a separation of 2800A. The obvious probable explanation of the anomaly of the lower curve is suggested by the fact that when the circuit inductance is higher, as in the upper curve, there is nothing peculiar in the distribution. One can postulate that the upper curve is approximately correct for both values of inductance. The anomalous peak at 2800A for the lower inductance can be made up of tests in which an arc occurred not at a separation of 2800A but at a much smaller separation, if only we imagine that the mound of metal thrown across between the electrodes could not solidify because of the magnitude of the current flowing through it. With the current sufficiently large the metal is prevented even from condensing to a liquid and the arc is therefore maintained. According to this theory a metal bridge is formed across the electrodes only after the current has decreased sufficiently. One must imagine that this effect, resulting in a spurious measure of electrode separation at arcing, occurs when the inductance is 0.10 μh but not when the inductance is 0.23 μh. For the former inductance the maximum arc current is 1.5 amp and the arc time is 0.020 μsec, and for the latter 1.0 amp and 0.030 μsec. Simple calculations indicate that these hypotheses are not unreasonable, but because the times are so short that a condition of steady heat flow does not exist no satisfactory calculations seem feasible.

From the distribution of Fig. 7 one concludes that the average separation at which an arc strikes between

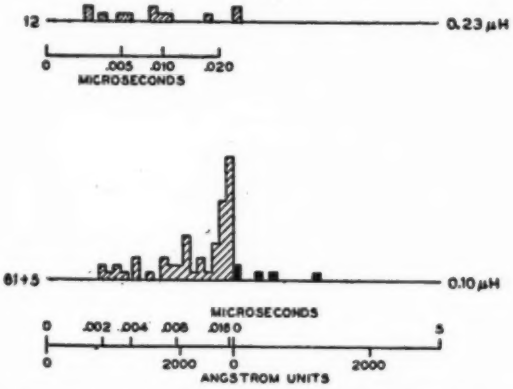


FIG. 7. The distribution of electrode separations at which an arc strikes between inactive silver electrodes.

TABLE I. Distance at which an arc strikes, and critical electric field.

	Reference 4 Table III	Reference 2 Table I		This paper	
Metal	Platinum	Silver	Silver	Silver	Silver
Surface condition	(?)	Active	Inactive	Active	Inactive
Voltage	48	35	35	40/45	35
Distance					
Uncorrected	9100A	6000A	1600A	2500A	—
Correction	—	1200A	1200A	2800A	1500A
Corrected	9100A	7200A	2800A	5300A	1500A
Gross Field volts/cm	0.53×10^6	0.49×10^6	1.2×10^6	0.8×10^6	2.3×10^6

approaching inactive silver electrodes is about 1500A corresponding to a gross electric field of about 2.3×10^6 volts/cm. The corresponding average separation of active electrodes of 2500A from Fig. 6 must clearly be corrected by the height of the mound of 2800A (from Eq. (1) with $\epsilon=1$) thrown up by the arc, giving 5300A and an average gross field of 0.8×10^6 volts/cm. These values agree extremely well with those reported earlier[11] for silver electrodes but the earlier separations were not corrected[12] for the mound height and after such correction there is a considerable discrepancy. The new values of average critical field are considerably higher than the old. The available values from the two sets of measurements are collected in Table I, together with still earlier data for platinum (reference 4, Table III). The spread in any group of measurements is obviously very great and perhaps the agreement between the different sets is as good as can be expected.

It is interesting to point out that in the earlier studies[2] with inactive silver electrodes we concerned outselves only with arcs which end in open circuits whereas in the present tests very few opens are found following arcs. The reason for finding many opens in the earlier tests and few in those reported here is to be found in the difference between the sizes of the condensers which are discharged in the two cases, 40 $\mu\mu$f in the earlier tests and 400 $\mu\mu$f in the present. These two capacities yield by Eq. (1) mounds of 1300A and 2800A, respectively; for the former one expects to find frequent opens, following an arc between inactive electrodes and for the latter very few opens, both of which predictions are in agreement with the facts.

For the present measurements the average critical electric field at which a short arc strikes between inactive silver electrodes is 2.3×10^6 volts/cm and between active silver electrodes 0.8×10^6 volts/cm. These are probably more reliable than earlier values.

The Formative Period of the Arc

Oscilloscopic traces of the voltage across active silver electrodes at the beginning of a short arc are repro-

[11] See Table I of reference 2.

[12] The correction was worked out in footnote 3 on page 1136 of reference 2, but was not applied because of the uncertainties involved.

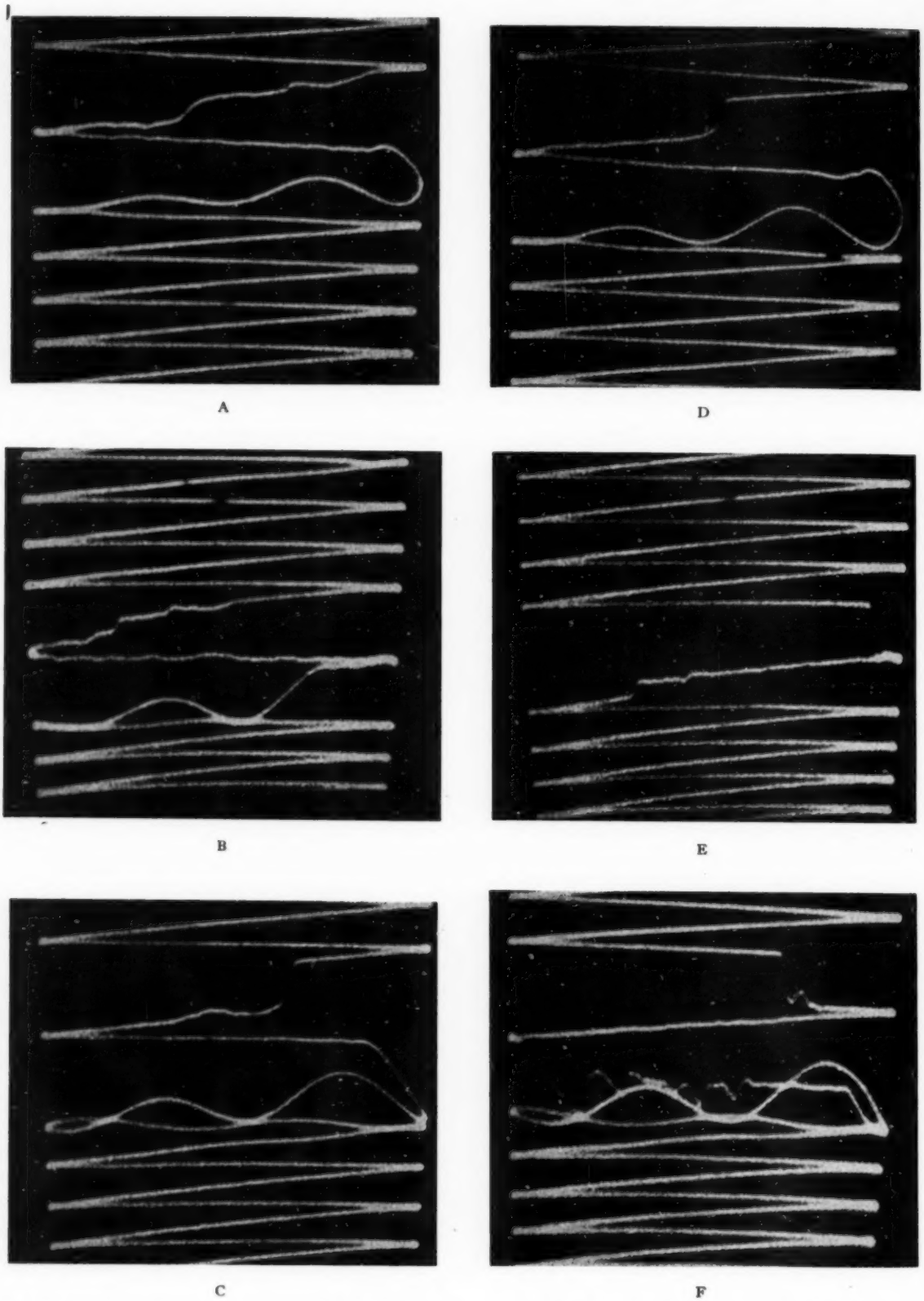
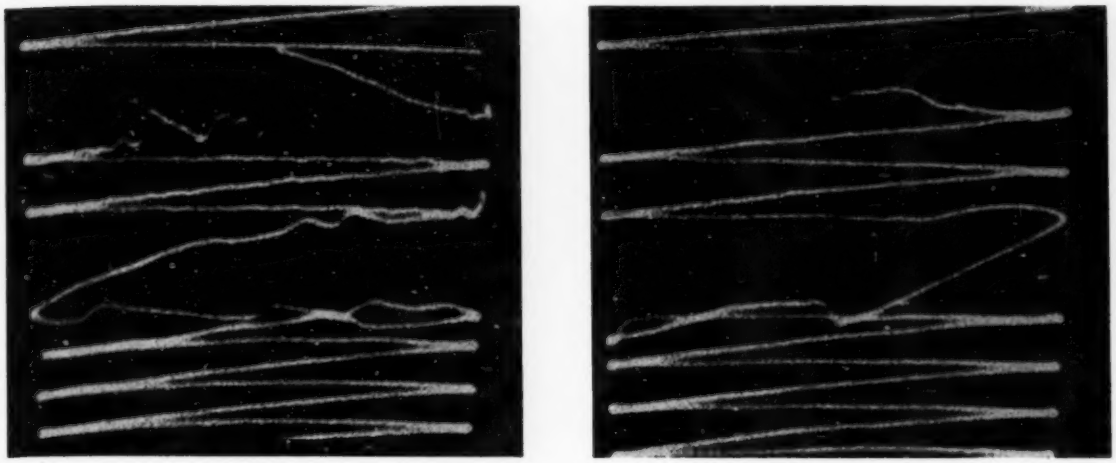


FIG. 8. Active silver electrodes discharging a 400 $\mu\mu$f condenser through 5.0 μh. (V_0=35 volts in A, B, and D and 40 volts in C, E, and F.) The calculated arc time of 1.4 trace lengths agrees well with the observed arcs (except for E).



A B

FIG. 9. Active silver electrodes discharging a 400 $\mu\mu$f condenser through 48 μh. ($V_0=40$ volts.) The calculated arc time is 4.0 trace lengths.

duced in Figs. 4, 8, and 9. The inductances in the circuits which produced these traces are, respectively, 0.10, 5.0, and 48 μh. The capacity which is discharged is 400 $\mu\mu$f. In each of the traces of Figs. 8 and 9 one sees a steady arc at a constant voltage, but the way in which the potential decreases before the steady value becomes established is highly variable.

One expects the duration of the arc to be somewhat less than the nominal value of $\pi(LC)^{\frac{1}{2}}$ sec, because each arc should be extinguished when the current through it reaches the minimum arc current i_0 rather than when the current becomes zero. If one neglects the deviation of the arc potential during the beginning of the arc from its normal value after it has become established, the arc time is calculated to be

$$t_0=\pi(LC)^{\frac{1}{2}}\{1-(1/\pi)\sin^{-1}[i_0/(V_0-v)(C/L)^{\frac{1}{2}}]\}.$$

For $i_0=0.03$ amp[11] this gives $t_0=\pi(LC)^{\frac{1}{2}}(0.963)=0.135$ μsec for $L=5$ μh, and $t_0=\pi(LC)^{\frac{1}{2}}(0.886)=0.386$ μsec for $L=48$ μh. These times are, respectively, 1.4 and 4.0 average sweeps across the screen (0.096 μsec for each average sweep), which agree well[13] with the arcs of Figs. 8 and 9, except for the single trace of Fig. 8(E) on which the contacts are observed to close before the arc is over. Early closure is rare for active silver electrodes, with very few traces resembling that of Fig. 8(E).

Among the traces of Fig. 8 there is a wide variation in the elapsed times from the beginning of an arc until the steady arc voltage is established, ranging from about 0.04 μsec for Figs. 8(A) and 8(B), to 0.004 μsec for Fig. 8(D), and to perhaps 0.002 or 0.001 μsec for Fig. 8(E). This great variability is typical of the photographs. The average for a representative group of 58 traces is 0.022 μsec.The trace of Fig. 8(C) represents a closure showing about the average time for the steady value to be established, with Figs. 8(A) and 8(B) closures requiring more than average time and 8(D), 8(E), and 8(F) less than average time. One might guess that the steady arc potential would be established more slowly for an arc between electrodes far apart than for an arc between electrodes close together. No such correlation has, however, been observed; as far as we can tell the arcs between electrodes relatively close together show about the same range of variation of the times for the steady voltage to become established as do arcs between electrodes at the greatest separations which have been noted.[14] This observation must be interpreted to mean that the time required for a silver ion to move from one electrode to the other is not the dominant factor in determining how quickly the steady arc potential becomes established.

For an inductance as large as 48 μh the most usual closure is one for which there are repeated discharges of the oscilloscope plate capacity before finally the 400 $\mu\mu$f capacity begins to discharge steadily in a sustained arc; the pattern of Fig. 9(A) is typical of many oscilloscope traces, but that of Fig. 9(B) in which the main arc is established without such successive small discharges is rather rare. The limiting inductance at which the two types of closure are equally frequent is between 5 μh and 48 μh, and perhaps in the neighborhood of 20 μh.

[13] In principle the minimum arc current i_0 can be determined from a measured value of t_0 by means of the above equation. This method has little merit.

[14] Many closures in the 5 μh circuit give oscilloscope traces much more complicated than those of Figs. 8(A)-(E). An example of such a complex trace is given in Fig. 8(F). On this trace the open circuit oscillation following the sustained arc is interrupted at least five times by arcs of short duration in the reverse direction, the last of these being the longest and lasting for about 0.025 μsec, after which the open oscillation is continued until its energy is dissipated in the 200 ohm resistance. The closure represented by the trace of Fig. 8(F) is more complicated than the average closure in the 5 μh circuit, but it is on the other hand less complex than a great many closures whose traces have been photographed. In many of these there are repeated discharges of the oscilloscope plate capacity each ending in an open oscillation, before the beginning of the principal arc. These have been described earlier as "arrested discharges."

Many traces have been recorded of closures for circuit inductances of 200 and 1000 μh. None of these exhibits a sustained arc; each is made up of repeated discharges of small local capacity, separated by recharges from the 400 $\mu\mu$f through the large inductance. It is one of these discharges of local capacity, or a series of them, followed by a considerable interval without discharge, which makes up what we have described earlier[16] as an "arrested discharge."

It seems intuitive that the time required for a steady arc to be established between active electrodes should vary with circuit inductance being shorter for small values than for large, simply because of the larger current when the inductance is small. Observations upon active electrodes seem to support this guess. For the lowest inductance of 0.10 μh the average time for a group of 36 representative closures has been measured to be about 0.011 μsec, for 5 μh the average mentioned above is 0.022 μsec, and for 48 μh an arc is usually established only after the lapse of a much longer time and after several discharges and recharges of the local capacity. Comparable measurements, which are necessarily very crude indeed, have been made upon the closure of inactive electrodes with a circuit inductance of 0.10 μh; the average of 47 such measurements is 0.007 μsec. Up to the present time we have had no success in drawing any fundamental understanding of the arc from these simple observations.

[16] Reference 4, pages 1105–1107.

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 5 MAY, 1952

Instabilities in the Smooth-Anode Cylindrical Magnetron*

LAWRENCE A. HARRIS†
University of Florida, Gainesville, Florida
(Received January 7, 1952)

A magnetically focused space-charge cloud of the type found in the smooth-anode magnetron is examined to determine whether a perturbation of the equilibrium condition will grow. A field analysis is carried out in which radial admittances are matched at the edge of the cloud. The solutions for the characteristic frequencies are complex, indicating that the disturbance grows with time. These instabilities occur at multiples of the Larmor frequency.

The solutions indicate that electrons oscillate with increasing swings, a process which absorbs energy. Instability therefore gives rise to some anode current even when the tube is "cut off" though rf power is not made available by the process.

INTRODUCTION

THE residual current conduction in the smooth-anode magnetron with the anode voltage below the critical value has long been a well-known but unexplained phenomenon. This current flow is considerably greater than that accounted for by thermal emission velocities and such considerations as mechanical misalignment of tube parts.[1] It has been generally supposed that the explanation of the excess current lies in the presence of some kind of oscillation in these tubes.

This hypothesis has been supported by scattered experimental observations of high frequency radiation[2] for such tubes and the back-heating of magnetron cathodes.[3] The present paper is an investigation of instabilities in space-charge clouds of the type believed to be found in magnetrons. The general analysis carried out for small perturbations on the steady-state conditions is completed for the special case of the smooth-anode magnetron with vanishingly small cathode.

The method of analysis followed is to derive the radial wave admittance for waves in the charge-free region of the tube. A similar quantity is derived from conditions in the interior of the space-charge cloud, and these two admittances are matched at the edge of the cloud.[4] The treatment of the dynamics of the electron cloud involving the use of the action function follows the methods of Buneman[5] and of MacFarlane and Hay.[6] The latter authors have analyzed the behavior of a slipping stream of electrons as encountered in the plane magnetron. The circular magnetron is considered here because of its two essential differences: It closes on itself imposing a periodicity condition on the waves, and the electron stream is not a slipping one.

In a sense the analysis is somewhat unrealistic. The usual steady-state conditions described by Brillouin[7,8]

* This work has been supported in part by the ONR.
† Now at the University of Minnesota, Minneapolis, Minnesota.
[1] Sarbacher and Edson, *Hyper and Ultra-High Frequency Engineering* (John Wiley & Sons, Inc., New York, 1943), Chapter 16, p. 552.
[2] G. H. Geick, T. R. No. 14, Stanford University, April 30, 1949.
[3] Young, Holmboe, and Walters, J. Appl. Phys. **21**, 1066 (1950).
[4] L. J. Chu and J. D. Jackson, Proc. Inst. Radio Engrs. **36**, 853 (1948).
[5] O. Buneman, Proc. Phys. Soc. (London) **B63**, 278 (1950).
[6] G. G. MacFarlane and H. G. Hay, Proc. Phys. Soc. (London) **B63**, 409 (1950).
[7] L. Brillouin, Phys. Rev. **60**, 385 (1941).
[8] L. A. Harris, Axially Symmetric Electron Beam and Magnetic Field Systems, T. R. No. 170, August 29, 1950, Research Lab. of Electronics, M.I.T.

are assumed to exist in the tube, and then this state is examined for stability. If it turns out to be a stable equilibrium condition, then the results obtained are valid and useful. On the other hand, if the Brillouin flow is found to be an unstable equilibrium, very little information is gained as to actual conditions in these tubes. It is difficult to conceive of an unstable equilibrium being established, and then breaking down as two separate processes; they would be expected to occur simultaneously. Moreover, if self-excited oscillations are present, the process must be a nonlinear one.

The following small-signal analysis which does indeed indicate the possibility of oscillations in the magnetron should be considered in the light of the previous remarks. The information gained is limited to the fact of instability and to an approximate prediction of the oscillation frequencies.

ADMITTANCE IN THE CHARGE-FREE REGION

The geometry of the magnetron and the coordinate system used are shown in Fig. 1. The anode of radius r_a is assumed to be a perfectly conducting circular cylinder. The electron cloud with outer radius r_e is assumed to exist in the Brillouin steady state. M.K.S. rationalized units are used.

Since any oscillation must depend on the interaction of electric fields and electrons it is assumed that the field responsible for the interaction is one with a circumferential component of electric field. Thus the waves considered are transverse magnetic (TM) when viewed as propagating around the tube but are transverse electric (TE) when considered as the more usual type of waveguide mode. It is difficult to take account of any axial variation in the electron clouds of these tubes, and experiments indicate that the tube length has little effect on the process, so axial variation is neglected in the charge-free region also.

The field components of a TE wave-guide mode may be expressed in terms of the longitudinal component of magnetic field H_z which must satisfy the Helmholtz equation

$$\frac{\partial^2 H_z}{\partial r^2}+\frac{1}{r}\frac{\partial H_z}{\partial r}+\frac{1}{r^2}\frac{\partial^2 H_z}{\partial \theta^2}+k^2 H_z=0, \tag{1}$$

where $k=\omega/c$ is the propagation constant of a plane wave in free space, ω is the angular frequency, and c is the velocity of light.

A solution to this equation is

$$H_z=[C_1 J_n(kr)+C_2 N_n(kr)]e^{-jn\theta}, \tag{2}$$

where C_1 and C_2 are constants and $J_n(kr)$ and $N_n(kr)$ are Bessel functions of the first and second kinds and order n. The circularity of the region in which this solution applies requires that n be an integer. Application of Maxwell's equations yields the expression for the electric field E_θ

$$E_\theta=(j\omega\mu/k)[C_1 J_n'(kr)+C_2 N_n'(kr)]e^{-jn\theta}, \tag{3}$$

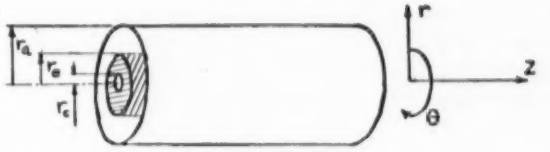


FIG. 1. A cylindrical smooth-anode magnetron operating at a voltage below cutoff.

where J_n' or N_n' represents the derivative of the Bessel function with respect to the argument kr.

Since E_θ vanishes at the anode surface

$$C_2=-C_1[J_n'(kr_a)/N_n'(kr_a)]. \tag{4}$$

The radial admittance Y_1 looking outward is defined as the ratio H_z/E_θ and this yields

$$Y_1=-\frac{jk}{\omega\mu}\left[\frac{J_n(kr)N_n'(kr_a)-J_n'(kr_a)N_n(kr)}{J_n'(kr)N_n'(kr_a)-J_n'(kr_a)N_n'(kr)}\right]. \tag{5}$$

Before this expression can be put to use another expression for the radial admittance governed by conditions inside the charge cloud must be developed. Both of these expressions are then evaluated at the cloud radius r_e and equated. The solution of the resulting equation yields the characteristic frequencies of oscillation.

To find the second radial admittance the dynamics of the electrons must be considered. A difficulty arising here is the two-dimensional character of the electron motion including both radial and circumferential velocities. The usual equations of motion which express acceleration in terms of the electric and magnetic fields acting on the electrons become rather involved, and it is found to be more profitable to formulate the problem in terms of an action function or velocity potential.

It has been shown by Gabor[9] and Buneman[5] that in vacuum tubes with no magnetic flux intersecting the emitting surface of the cathode, the use of an action function is possible, since the generalized momentum[8] has no curl. This momentum may be expressed as the gradient of an action function F

$$p=mv+eA=\nabla(mF), \tag{6}$$

where p is the generalized momentum, v is the vector velocity, and A is the vector magnetic potential. The charge e is the algebraic electron charge, a negative number, and m is the electron mass. The use of this action function in an expression derived from the Lorentz force equation results in Bernoulli's equation[5,6]

$$v^2+2(\partial F/\partial t)=-2(e/m)\varphi, \tag{7}$$

where φ is the electric potential.

Elimination of v in Eq. (7) by means of Eq. (6) yields

$$(\partial/\partial t)(mF)+(m/2)[\nabla F-(e/m)A]^2=-e\varphi. \tag{8}$$

At this point it is postulated that A has a θ component only and that all the pertinent quantities consist of a

[9] D. Gabor, Proc. Inst. Radio Engrs. 33, 792 (1945).

steady component plus a small perturbation characterized by the wave factor $e^{j\psi}$ where $\psi=\omega t-n\theta$ and n is chosen to conform with the wave solution 2 and 3. Consequently the action function is given by

$$F=F_0(r,\theta)+F_1(r)e^{j\psi}. \tag{9}$$

This quantity is substituted into Eq. (8) and only the linear terms in $e^{j\psi}$ are retained, the higher order terms being considered negligible. The resulting equation is

$$mj\omega F_1 e^{j\psi}+\frac{m}{2}\left\{\left(\frac{\partial F_0}{\partial r}\right)^2+2\frac{\partial F_0}{\partial r}\frac{\partial F_1}{\partial r}e^{j\psi}+\left[\frac{1}{r}\frac{\partial F_0}{\partial\theta}-\frac{e}{m}A_\theta\right]^2-2\frac{n}{r}F_1\left(\frac{1}{r}\frac{\partial F_0}{\partial\theta}-\frac{e}{m}A_\theta\right)e^{j\psi}\right\}=-e\varphi. \tag{10}$$

When the perturbation F_1 is absent, the condition of the electron cloud is that described by the Brillouin steady state. In this case there is no radial velocity and consequently

$$\partial F_0/\partial r=0, \tag{11}$$

while the steady component φ_0 of the potential is given by

$$m/2[(1/r)(\partial F_0/\partial\theta)-(e/m)A_\theta]^2=-e\varphi_0. \tag{12}$$

It follows from these equations that the perturbation term in the potential expressions is given by

$$\varphi_1=-j\omega(m/e)[1-(nv_{0\theta}/\omega r]F_1, \tag{13}$$

where the total potential is

$$\varphi=\varphi_0+\varphi_1e^{j\psi} \tag{14}$$

and $v_{0\theta}$ is the steady circumferential electron velocity. The expression (14) is now substituted into Poisson's equation

$$\frac{\partial^2\varphi}{\partial r^2}+\frac{1}{r}\frac{\partial\varphi}{\partial r}+\frac{1}{r^2}\frac{\partial^2\varphi}{\partial\theta^2}=-\frac{(\rho_0+\rho_1e^{j\psi})}{\epsilon}, \tag{15}$$

and the time-invariant terms are equated to the steady component of charge density, leaving

$$(d^2\varphi_1/dr^2)+(1/r)(d\varphi_1/dr)-(n^2/r^2)\varphi_1=-\rho_1/\epsilon. \tag{16}$$

Substituting (13) into (16), an expression for ρ_1, in terms of F_1 and $v_{0\theta}$, is obtained:

$$\rho_1=j\omega\epsilon\frac{m}{e}\left[\eta\frac{d^2F_1}{dr^2}+\left(2\frac{d\eta}{dr}+\frac{\eta}{r}\right)\frac{dF_1}{dr}+\left(\frac{d^2\eta}{dr^2}+\frac{1}{r}\frac{d\eta}{dr}-\frac{n^2\eta}{r^2}\right)F_1\right], \tag{17}$$

where

$$\eta=1-(nv_{0\theta}/\omega r). \tag{18}$$

Another expression for ρ_1 is now obtained from the continuity equation and equated to (17) to yield a differential equation which must be satisfied in the magnetron. First it is necessary to calculate the components of current density due to the electron velocities. The components of electron velocity are determined from the vector equation (6). The radial current density is

$$J_r=(\rho_0+\rho_1e^{j\psi})(v_{0r}+v_{1r}e^{j\psi}), \tag{19}$$

which on application of (6) and neglecting higher order terms, simplifies to

$$J_r=\rho_0(dF_1/dr)e^{j\psi}, \tag{20}$$

and the angular current density J_θ is

$$J_\theta=(\rho_0+\rho_1e^{j\psi})(v_{0\theta}+v_{1\theta}e^{j\psi}), \tag{21}$$

which in turn reduces to

$$J_\theta=\rho_0v_{0\theta}-j\rho_0(n/r)F_1e^{j\psi}+v_{0\theta}\rho_1e^{j\psi}. \tag{22}$$

On substituting these expressions into the continuity equation

$$\frac{\partial J_r}{\partial r}+\frac{1}{r}J_r+\frac{1}{r}\frac{\partial J_\theta}{\partial\theta}+\frac{\partial\rho}{\partial t}=0 \tag{23}$$

and solving for ρ_1 the following is obtained:

$$\rho_1=\left[\rho_0\frac{\partial^2F_1}{dr^2}+\left(\frac{\rho_0}{r}+\frac{d\rho_0}{dr}\right)\frac{dF_1}{dr}-\rho_0\frac{n^2}{r^2}F_1\right]\Big/-j\omega\eta. \tag{24}$$

Equating Eqs. (17) and (24) for ρ_1 and simplifying yields the differential equation for $F_1(r)$

$$\left(\frac{d^2F_1}{dr^2}+\frac{1}{r}\frac{dF_1}{dr}-\frac{n^2}{r^2}F_1\right)\left(\frac{\omega_p^2}{\omega^2}-\eta^2\right)+\frac{dF_1}{dr}\frac{d}{dr}\left(\frac{\omega_p^2}{\omega^2}-\eta^2\right)-F_1\eta\left(\frac{d^2\eta}{dr^2}+\frac{1}{r}\frac{d\eta}{dr}\right)=0, \tag{25}$$

where

$$\omega_p^2=(e/m)(\rho_0/\epsilon) \tag{26}$$

defines the plasma frequency ω_p. The solution to (25) depends on the particular type of tube being considered. In the form shown it applies equally well to any tube with a uniform axial magnetic field and a cathode not penetrated by magnetic flux. The parameters appearing in the equation depend on the dc focusing conditions; in particular the form of η will vary with the physical arrangement.

In deriving the radial admittance for waves in the space-charge cloud it will be necessary to solve the differential equation. The reason for this may be appreciated after developing an expression for this admittance.

CALCULATION OF RADIAL ADMITTANCE FROM CONDITIONS IN THE ELECTRON CLOUD

In determining the admittance outside the charge region the field components H_z and E_θ were found and their ratio taken. The same two components are to be found inside the electron cloud. The sum of the radial convection and displacement currents is used to find H_z by means of the Maxwell equation

$$\nabla\times H=J+\epsilon(\partial E/\partial t), \tag{27}$$

and the angular component of electric field E_θ is readily evaluated from the potential φ. It is not the magnetic field inside the cloud that is required here though, but that just outside the cloud. The distortion of the electron cloud results in an effective surface current density which produces a discontinuity in the magnetic field. With this discontinuity taken into account, it is possible to find the radial admittance which must be matched to that expressed in Eq. (5).

As a result of the interaction of the fields with the electrons, the cylindrical surface of the cloud is distorted, the humps or depressions appearing to travel around it as waves. It is assumed that the equation of this surface is

$$S=r-(r_e+\delta e^{j\psi})=0, \tag{28}$$

where $\delta\ll r_e$, the unperturbed radius.

Any electron moving in the surface of the cloud must have a velocity tangential to it, so

$$\frac{dS}{dt}=\frac{\partial S}{\partial t}+\frac{\partial S}{\partial r}v_r+\frac{\partial S}{\partial\theta}\frac{v_\theta}{r}=0. \tag{29}$$

Substituting (28) into (29) and solving for δ gives

$$\delta=\frac{v_r}{j\omega e^{j\psi}\eta}=-\frac{j(\partial F_1/\partial r)}{\omega\eta}. \tag{30}$$

The radial current densities at the radius r are derived from (20) and (13):

$$J_r+\epsilon\frac{\partial E_r}{\partial t}=\left(\rho_0\frac{\partial F_1}{\partial r}-j\omega\epsilon\frac{\partial\varphi_1}{\partial r}\right)e^{j\psi}. \tag{31}$$

This quantity according to Eq. (27) must be $(1/r)(\partial H_z/\partial\theta)$ since H_θ is assumed to be zero. But if

$$H_z=H_0+H_1e^{j\psi}, \tag{32}$$

then

$$\partial H_z/\partial\theta=-jnH_1e^{j\psi} \tag{33}$$

or

$$H_1=\frac{jr\rho_0}{n}\frac{\partial F_1}{\partial r}+\frac{\omega\epsilon r}{n}\frac{\partial\varphi_1}{\partial r}. \tag{34}$$

This is the perturbation on the magnetic field; i.e., that part of the magnetic field due to the wave in the cloud. Between this radius r_e and the true surface at $r_e+\delta e^{j\psi}$ there is an angular convection current (Fig. 2) per unit

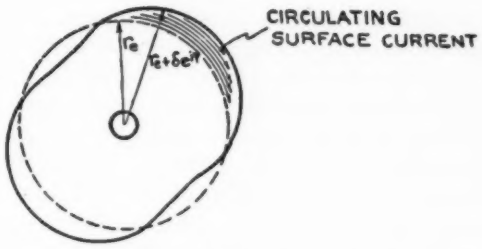


FIG. 2. Deformation of the charge cloud and the equivalent surface current.

length along the axis of approximately

$$J_{s\theta}=\rho_0v_{0\theta}\delta e^{j\psi}=-\frac{j\rho_0v_{0\theta}(\partial F_1/\partial r)}{\omega\eta}e^{j\psi}. \tag{35}$$

This current is considered to be compressed into an infinitesimal layer at r_e, thus appearing as a surface current density of the value given in (35). This procedure furnishes the cylindrical coordinate surface at which boundary conditions may be applied.

The apparent surface current density is responsible for a discontinuity in the magnetic field H_z, which gives a field just outside the cloud of

$$H_z(r_e+\delta e^{j\psi})=H_0+\left[\frac{j\rho_0v_{0\theta}(dF_1/dr)}{\omega\eta}+j\rho_0\frac{r}{n}\frac{dF_1}{dr}+\frac{\omega\epsilon r}{n}\frac{\partial\varphi_1}{\partial r}\right]e^{j\psi}, \tag{36}$$

and the part of this field due to the wave is, on application of (13),

$$H_1(r_e+\delta e^{j\psi})=\rho_0\frac{\partial F_1}{dr}\left[\frac{jv_{0\theta}}{\omega\eta}+j\frac{r}{n}-j\frac{\omega^2\epsilon r}{n\rho_0}\frac{m}{e}\eta\right]-j\frac{\omega^2\epsilon r}{n}\frac{m}{e}F_1\frac{d\eta}{dr}. \tag{37}$$

The electric field E_θ is easily derived from the potential φ which, because of (13) yields

$$E_\theta=-(1/r)(\partial\varphi/\partial\theta)=-(n/r)(m/e)\omega\eta F_1e^{j\psi}. \tag{38}$$

Therefore the outward radial admittance just outside the space-charge cloud, but derived from conditions inside, is given by

$$Y_2=\frac{H_1e^{j\psi}}{E_\theta}=\frac{\rho_0\dfrac{dF_1}{dr}\left[\dfrac{jv_{0\theta}}{\omega\eta}+j\dfrac{r}{n}-j\dfrac{\omega^2\epsilon r}{n\rho_0}\dfrac{m}{e}\eta\right]-j\dfrac{\omega^2\epsilon r}{n}\dfrac{m}{e}F_1\dfrac{d\eta}{dr}}{-(nm/re)\omega\eta F_1}. \tag{39}$$

This is the quantity that must be equated to the admittance as given by (5). It is apparent from (39) that before Y_2 can be properly evaluated the quantity $(1/F_1)(dF_1/dr)$ must be found by solving the differential

equation (25). In order to do so, the parameters η and ω_p^2 must be determined from the steady-state (unperturbed) conditions in the tube. The particular case of the smooth-anode magnetron with vanishingly small cathode lends itself to a ready solution.

The equilibrium conditions for the magnetron are well known. It is interesting to note that when the cathode is very small, these conditions approach those in the solid magnetically focused electron beam.[8] It is as though the cathode were shrunk in diameter and then pulled axially out of the tube and out of the magnetic field where it may be expanded again into the flat cathode of the electron gun. The parameters for the magnetron or the solid beam are

$$\omega_p^2=2\omega_H^2 \tag{40}$$

and

$$\eta=1+(n\omega_H/\omega), \tag{41}$$

where

$$\omega_H=(e/m)(B/2). \tag{42}$$

Consequently

$$d\omega_p^2/dr=0 \tag{43}$$

and

$$d\eta/dr=0 \tag{44}$$

so the differential equation reduces to

$$\frac{d^2F_1}{dr^2}+\frac{1}{r}\frac{dF_1}{dr}-\frac{n^2}{r^2}F_1=0. \tag{45}$$

The solution to this equation is

$$F_1=A_1r^n+A_2r^{-n}, \tag{46}$$

where A_1 and A_2 are constants.

This solution cannot become infinite at the axis in the case of the solid beam, and in the magnetron similar conditions are expected even though a small cathode is present. Therefore only the term with the positive exponent is retained in the solution. The required ratio appearing in (39) is then

$$(1/F_1)(dF_1/dr)=n/r. \tag{47}$$

With the parameter values expressed in (40) and (41) and the ratio n/r substituted into (39), the radial admittance Y_2 becomes

$$Y_2=j\omega\epsilon(r/n)[1-2\omega_H^2/(\omega+n\omega_H)^2]. \tag{48}$$

Finally the transcendental equation determining the characteristic frequencies of the smooth-anode magnetron is

$$j\omega\epsilon\frac{r_c}{n}\left[1-\frac{2\omega_H^2}{(\omega+n\omega_H)^2}\right]=-\frac{jk}{\omega\mu}\left[\frac{J_n(kr_c)N_n'(kr_a)-J_n'(kr_a)N_n(kr_c)}{J_n'(kr_c)N_n'(kr_a)-J_n'(kr_a)N_n'(kr_c)}\right]. \tag{49}$$

SOLUTION OF THE TRANSCENDENTAL EQUATION (49)

The allowed integral values of n correspond to different modes of oscillation, n representing the number of full waves around the circumference of the tube. If the electron cloud is unstable the frequency ω characteristic of a particular mode must be complex, the imaginary part of ω giving the rate of exponential growth, or decay of the assumed perturbation.

A general solution of (49) for complex values of ω is difficult and laborious especially for the higher values of n because of the lack of tabulated Bessel functions of complex argument. If it is assumed, however, that the dimensions of the tube and the magnetic field intensity are limited in value so the arguments of the Bessel functions are much less than 1, then the equation becomes tractible. [If the anode radius is 0.635 cm (1/4 inch) and it is assumed that for $n=1$, $|\omega|$ is of the order of $|\omega_H|$, then the magnetic field is limited to about 500 gauss if kr_a is not to exceed 0.1.]

For these small arguments the Bessel functions are approximated by the first term in their series expansions. Thus

$$J_n(x)\approx(1/n!)(x/2)^n$$

and

$$N_n(x)\approx-[(n-1)!/\pi](2/x)^n \quad n\neq 0 \tag{50}$$
$$J_n'(x)=[1/(n-1)!(x^{n-1}/2^n)$$

and

$$N_n'(x)=n!/\pi(2^n/x^{n+1}). \tag{51}$$

With these substitutions the right-hand side of Eq. (49) is simplified, and therefore

$$1-[2\omega_H^2/(\omega+n\omega_H)^2]=-[(r_c^{2n}+r_a^{2n})/(r_c^{2n}-r_a^{2n})], \tag{52}$$

which is easily solved for ω. The result is

$$\omega=-n\omega_H\pm[2\omega_H^2/1-g_1]^{\frac{1}{2}}, \tag{53}$$

where g_1 is a geometrical factor given by

$$g_1=\frac{(r_a/r_c)^{2n}+1}{(r_a/r_c)^{2n}-1}. \tag{54}$$

It is significant that g_1 is positive and greater than 1. The second term of the solution is therefore imaginary, indicating that either a growing or decaying exponential amplitude is possible. The general solution for this system, because it is linear for small perturbations, consists of a linear combination of both characteristic frequencies. The part of the solution giving rise to a decreasing wave cannot cause any nonlinearities in the system and, therefore, does not prevent the existence of the growing wave. Thus in any case, it may be concluded that the growing wave is present and the space-charge cloud is unstable.

The frequencies of oscillation occur at integral multiples of the Larmor frequency ω_H, and curiously the growth factor or gain appears to increase with the mode

number n. It should be recalled, however, that as the mode number increases, the range of values of magnetic field for which the approximations (50) are valid decreases considerably. Accordingly even for moderate magnetic fields the behavior of the higher mode oscillations must be calculated using better approximations to the Bessel functions than have been used here. Care must be taken too that inappropriate values of r_a and r_s are not applied in any particular case.

The geometrical factor g approaches 1 when the ratio r_a/r_s becomes large, and the gain increases. This should not be taken to mean that large outputs can be derived from a very thin beam. The only meaning to be attached to these results is the instability of the space-charge clouds or beams being considered. Once the equilibrium conditions are departed from appreciably, the linear analysis no longer applies. In the case of an extremely thin beam, saturation would probably occur very quickly; i.e., the equilibrium condition would break down rapidly as indicated by the large imaginary part of ω.

It will be noted that when the radius of the electron cloud approaches the anode radius g_1 becomes infinite and the gain falls to zero. This conclusion is to be expected as the smooth conducting anode wall cannot support any tangential electric field, hence there can be no interaction.

All of the above results apply for integral values of n greater than zero. When n is zero there is no circumferential electric field E_θ as indicated by the infinite value of Y_2 expressed in Eq. (48).

As might have been expected, the solution (53) indicates that the wave and the electrons rotate about the axis in synchronism. This may be seen by evaluating the parameter η given in Eqs. (18) and (42). The real part of η is zero, or the angular velocity $[\mathrm{Re}(\omega)/n]$ of the wave is equal to that of the electrons v_{00}/r.

CONCLUSIONS

It has been demonstrated that a particular type of magnetically focused space-charge cloud is an unstable configuration. While a number of approximations have been made, it seems reasonable to expect that more exact solutions to the problem would yield qualitatively similar results.

These instabilities are of a rather special sort as they are due to the effects of high frequency waves in the tube structure. The electron orbits are stable in the sense that if any one electron is displaced from its equilibrium path, it returns to that path or at least oscillates about it. The entire cloud is unstable, however, in the sense that such a disturbance affects the rest of the cloud, and the interaction between all the electrons and the fields in the tube tends to build up the process, thereby establishing self-excited oscillations.

Although these oscillations are not quite the same as the more familiar cyclotron oscillations, the paths followed by individual electrons have much the same appearance. They are circular motions superimposed on the Larmor drift and growing in radius as the amplitude of the disturbance increases.

The process discussed in this paper is to be distinguished also from that taking place in the more usual type of traveling-wave magnetron. In the latter case electrons drift toward the anode as they lose energy to the wave, but the amplitude of the oscillatory part of their motion does not change appreciably. This magnetron process has not been considered here, though it undoubtedly occurs in the type of tubes that have been discussed.

The ultimate behavior of electron clouds as it would be observed experimentally is difficult to predict from the analysis presented here. While the characteristic frequencies may be found approximately, it is expected that any physical oscillation would have an amplitude bringing it beyond the scope of the linear analysis.

There appears to be some doubt as to whether rf energy could ever be withdrawn from a tube oscillating by this mechanism. The increasing swing of the electrons means an increase in their kinetic energy at the expense of the wave. The decreasing swings also indicated by the solution for ω are not significant here because small oscillations are assumed to start with. It might be that the best way to test this theory experimentally is to measure the characteristics of these clouds as power absorbing devices rather than as generators.

The mechanism therefore appears to be one which is capable of extracting energy from the dc supply of the tube and converting it to relatively useless kinetic energy of the electrons with no means of providing an rf output.

The result is a flow of anode current beyond cutoff with very few other outward manifestations of what is going on in the tube.

ACKNOWLEDGMENT

The author is grateful to Dr. W. W. Harman for suggesting the problem and for many stimulating discussions on it.

The Effect of Velocity Distribution in a Modulated Electron Stream

DEAN A. WATKINS
Electron-Tube Laboratories, Hughes Aircraft Company, Culver City, California
(Received January 24, 1952)

A method of solving electron beam problems is described which takes into account the thermal velocity spread. The method is based on the Liouville theorem. Two such problems are solved by means of a power series. In the case of small-signal velocity modulation of a drifting stream, the velocity spread distorts the shape of the standing wave of current which is produced, in addition to lengthening the plasma wavelength and the electronic wavelength. For a drifting stream with full shot noise in each velocity class at the input, a standing wave of noise current is produced with deep minima. The size of the first minimum is found to increase as the square of frequency and with increasing velocity spread. That full uncorrelated shot noise can produce a standing wave of convection current is somewhat surprising. The present analysis demonstrates that space charge is the important factor in determining the behavior of noise currents in electron streams even at high frequencies, and that velocity spread plays a less important role.

INTRODUCTION

THE behavior of beam-type, high frequency tubes has been analyzed in considerable detail by a number of workers[1–4] on the assumption that all the electrons in the stream in a very small region have the same velocity, thus neglecting the thermal velocity distribution. Analyses of signal and noise in diodes, triodes, tetrodes, and pentodes have been carried out by means of a set of equations derived by Llewellyn which also depend upon the assumption of a single-valued velocity for the electron stream.[5–7] Such analyses have been quite successful in predicting observed results in most cases. Where large transit angles are involved and when applied to noise problems, serious doubts arise regarding the validity of this approximation.

The effect of the thermal velocity spread upon noise and signal has been calculated[8] neglecting space charge. Although the results of these calculations may apply to cases where the charge density is so low that space-charge forces are nonexistent, in most practical problems the effect of space charge appears to be an important factor.

It is the purpose of the present paper to describe a method of calculation which takes into account the thermal spread in velocity and the space-charge forces. The effect of the thermal spread on signal and noise in a drifting stream is calculated by means of a power series. Although the calculation is lengthy, the method is straightforward and the results can be expressed in simple form.

FORMULATION

Consider an electron stream moving parallel to the z-axis. The stream is assumed to be broad in the direction perpendicular to the z axis so that variations in the transverse direction may be neglected. The velocity coordinate is $u=dz/dt$. Thus, the electron stream in our model moves in a two-dimensional phase space with coordinates z and u. The "location" of the electron stream in this two-dimensional space is described by a distribution function $F(z, u, t)$ such that $[F(z, u, t)/e]\,dzdu$ is the number of electrons per unit cross section between z and $z+dz$ with velocity between u and $u+du$ at the time t. The magnitude of the electronic charge is e. MKS units are used throughout.

According to the Liouville theorem of statistical mechanics, the distribution function $F(z, u, t)$ has the property that its total derivative with respect to time is zero. Thus

$$dF/dt=0. \tag{1}$$

In terms of partial derivatives, this can be written

$$\frac{\partial F(z,u,t)}{\partial t}+u\frac{\partial F(z,u,t)}{\partial z}+\frac{du}{dt}\frac{\partial F(z,u,t)}{\partial u}=0. \tag{2}$$

This is the Boltzmann transport equation neglecting collisions. Although the actual effect of the individual encounters or collisions between electrons is not known due to the failure of the collision integrals to converge, estimates of collision frequency for an electron gas worked out by Vlasov[9] indicate that the average number of longitudinal collisions per electron during its flight through an ordinary vacuum tube is of the order of one percent or less.

The acceleration, du/dt, is related to the electric field, $E(z, t)$, through the force equation which for nonrelativistic electron velocities and no magnetic field is

$$du/dt=-\eta E(z, t) \tag{3}$$

where η is the magnitude of the electronic charge-to-mass ratio.

The convection current density, $i(z, t)$, may be obtained by integrating the distribution function as

[1] W. C. Hahn, G. E. Rev. **42**, 258 (1939).
[2] S. Ramo, Phys. Rev. **56**, 256 (1939).
[3] J. R. Pierce, *Traveling-Wave Tubes* (D. Van Nostrand Company, Inc., New York, 1950).
[4] D. A. Watkins, Proc. Inst. Radio Engrs. **40**, 65 (1952).
[5] A. J. Rack, Bell Syst. Tech. J. **17**, 592 (1938).
[6] F. B. Llewellyn and L. C. Peterson, Proc. Inst. Radio Engrs. **32**, 133 (1944).
[7] L. C. Peterson, Proc. Inst. Radio Engrs. **35**, 1264 (1947).
[8] D. K. C. MacDonald, Phil. Mag. **40**, 561 (1949).
[9] A. Vlasov, J. of Phys. **9**, 25 (1945).

follows:

$$i(z, t)=-\int_{-\infty}^{\infty} uF(z, u, t)du. \tag{4}$$

The total current density $I(t)$, which at a given time is the same for all values of z, is the sum of the convection current density and the displacement current density. Thus

$$I(t)=i(z, t)+\epsilon_0[\partial E(z, t)/\partial t] \tag{5}$$

where ϵ_0 is the dielectric constant of vacuum.

We now separate all quantities into their dc and first-order, time-dependent parts as follows:

$$\begin{aligned} i(z, t)&=-I_0+i_1(z)e^{j\omega t}\\ I(t)&=-I_0+I_1e^{j\omega t}\\ E(z, t)&=E_0(z)+E_1(z)e^{j\omega t}\\ F(z, u, t)&=F_0(z, u)+F_1(z, u)e^{j\omega t}. \end{aligned} \tag{6}$$

Time dependence of the form $e^{j\omega t}$ has been assumed since we are interested in steady-state phenomena. Using the relations (6) and (3) in (2) and (5), we obtain for the dc quantities

$$\begin{aligned} u\frac{\partial F_0(z, u)}{\partial z}-\eta E_0(z)\frac{\partial F_0(z, u)}{\partial u}&=0\\ -I_0=-\int_{-\infty}^{\infty} uF_0(z, u)du& \end{aligned} \tag{7}$$

and for the ac quantities

$$j\omega F_1(z, u)+u\frac{\partial F_1(z, u)}{\partial z}-\eta E_1(z)\frac{\partial F_0(z, u)}{\partial u}-\eta E_0(z)\frac{\partial F_1(z, u)}{\partial u}=0 \tag{8}$$

$$I_1=-\int_{-\infty}^{\infty} uF_1(z, u)du+\epsilon_0 j\omega E_1(z) \tag{9}$$

where products of ac quantities have been neglected as in the usual small-signal analysis. The use of (9) to eliminate E_1 from (8) leads to

$$j\omega F_1(z, u)+u\frac{\partial F_1(z, u)}{\partial z}-\frac{\eta}{\epsilon_0 j\omega}\frac{\partial F_0(z, u)}{\partial u}\left[I_1+\int_{-\infty}^{\infty} uF_1(z, u)du\right]-\eta E_0(z)\frac{\partial F_1(z, u)}{\partial u}=0. \tag{10}$$

Equation (10) completely specifies the behavior of the system. Given the dc electric field $E_0(z)$, the dc distribution function $F_0(z, u)$, and the total ac current density I_1, it is possible, at least in principle, to solve for $F_1(z, u)$ subject to a single boundary condition. Such a boundary condition might be a knowledge of $F_1(z, u)$ at some particular value of z. In a diode, the total ac current density I_1, which may result from the application of an ac voltage across the diode or from the flow of induced current in an external impedance, must be determined from external parameters. If the diode is considered to be open-circuited for ac, the total ac current density is zero. A study of the Llewellyn[6] equations indicates that there is a large class of problems for which setting the total ac current equal to zero will give the correct result for the ac quantities in the stream. An example of this is the diode with large transit angle (greater than about 10 radians) for any degree of space charge. In both the problems treated here the total ac current will be assumed to be zero.

Once the ac distribution function has been determined, the ac convection current density can be found from

$$i_1(z)=-\int_{-\infty}^{\infty} uF_1(z, u)du. \tag{11}$$

For an electron stream formed by emission from a thermionic cathode, the dc distribution function is given by the half-Maxwellian distribution[5] which may be written

$$F_0(z, u)=2\alpha I_0 s(u-u_s)\exp[-\alpha(u^2-u_s^2)], \tag{12}$$

where

$$\alpha=m/2KT_c, \tag{13}$$

and $s(u-u_s)$ is the unit step function defined by

$$\begin{aligned} s(u-u_s)&=0 \quad u<u_s\\ s(u-u_s)&=1 \quad u>u_s. \end{aligned} \tag{14}$$

The velocity of the slowest electrons in the stream is u_s given by

$$u_s=u_s(z)=[2\eta V_0(z)]^{\frac{1}{2}}, \tag{15}$$

where $V_0(z)$ is the dc potential at z with respect to that of the potential minimum in front of the cathode if the emission is space-charge limited or, of the cathode itself, if the emission is temperature limited. m is electronic mass, K is Boltzmann's constant, and T_c is the cathode temperature in degrees Kelvin.

The derivative of the dc distribution function with with respect to u appearing in Eq. (10) is given by

$$\frac{\partial F_0}{\partial u}=2\alpha I_0\{\delta(u-u_s)-s(u-u_s)2\alpha u\exp[-\alpha(u^2-u_s^2)]\}, \tag{16}$$

where $\delta(u-u_s)$ is the Dirac delta-function defined by

$$\begin{aligned} \delta(u-u_s)&=0 \quad u\neq u_s\\ \int_{-\infty}^{\infty}\delta(u-u_s)f(z, u)du&=f(z, u_s) \end{aligned} \tag{17}$$

where $f(z, u)$ is an arbitrary function of z and u.

DRIFTING STREAM WITH SMALL-SIGNAL VELOCITY MODULATION

For the case of a drifting stream with dc space charge neutralized, the dc electric field $E_0(z)$ is zero and Eq. (10) becomes

$$j\omega F_1(z,u)+u\frac{\partial F_1(z,u)}{\partial z}-\frac{\eta}{\epsilon_0 j\omega}\frac{\partial F_0(z,u)}{\partial u}\int_{-\infty}^{\infty}uF_1(z,u)du=0 \quad (18)$$

for $I_1=0$. Now assume that (18) has a solution of the form

$$F_1(z,u)=e^{-j(\omega z/u_s)}\sum_{n=0}^{\infty}f_n(u)z^n. \quad (19)$$

Use of (19) in (18) leads to the recursion formula

$$f_n=\left[\frac{\eta}{un\epsilon_0 j\omega}\frac{\partial F_0}{\partial u}\int_{-\infty}^{\infty}uf_{n-1}du+j\left(\frac{\omega}{u_s}-\frac{\omega}{u}\right)\frac{f_{n-1}}{n}\right]. \quad (20)$$

The required boundary condition is the specification of $f_0(u)$ which is given by (see Appendix I)

$$f_0(u)=F_1(0,u)=(-\eta V_1/u)(\partial F_0/\partial u) \quad (21)$$

where V_1 is the amplitude of the small ac modulating voltage applied across a gap at $z=0$.

From (11) and (19)

$$i_1(z)=-e^{-j(\omega z/u_s)}\sum_{n=0}^{\infty}z^n\int_{-\infty}^{\infty}uf_n(u)du. \quad (22)$$

Let

$$a_n=j\frac{u_s}{\eta V_1}\frac{u_s}{I_0}\frac{\omega_p}{\omega}\int_{-\infty}^{\infty}uf_n(u)du. \quad (23)$$

Then

$$i_1(z)=j\frac{\eta V_1}{u_s}\frac{I_0}{u_s}\frac{\omega}{\omega_p}e^{-j(\omega z/u_s)}\sum_{n=0}^{\infty}a_nz^n. \quad (24)$$

The first eight coefficients of (24) are given in Appendix II. In calculating the a_n's, it is convenient to define the following functions of σ, where $\sigma=eV_0/KT_c$ is a measure of the thermal velocity spread.

$$g_n=\frac{-u_s^{n+2}}{I_0}\int_{-\infty}^{\infty}\left(\frac{1}{u_s}-\frac{1}{u}\right)^n\frac{\partial F_0(u)}{\partial u}du. \quad (25)$$

The asymptotic forms of the g_n's are given in Appendix II out to $1/\sigma^3$.

If the asymptotic forms of the g_n's are put into the coefficients of (24) and terms higher than $1/\sigma$ are omitted, the series is found to be that of

$$i_1(z)\simeq -j\left(\frac{\eta V_1}{u_s}\right)\frac{I_0}{u_s}\left(1-\frac{3}{4\sigma}\right)\frac{\omega}{\omega_p}\times\sin\left[\left(1-\frac{3}{4\sigma}\right)\frac{\omega_p z}{u_s}\right]e^{-j[1-(1/2\sigma)](\omega z/u_s)}, \quad (26)$$

where

$$\omega_p=(\eta I_0/\epsilon_0 u_s)^{\frac{1}{2}}. \quad (27)$$

ω_p is the plasma radian frequency which would exist if there were no thermal spread.

As $\sigma\to\infty$, Eq. (26) approaches the usual result for a space-charge standing wave obtained from single-valued velocity derivations.[1,2] It is seen that the first-order effect of the thermal spread is to reduce the amplitude, lengthen the optimum bunching distance, and lengthen the electronic wavelength.

To determine more accurately the effect of the thermal spread on the amplitude of the convection current, the series for $i_1(z)$ in (24) is multiplied by its complex conjugate which leads to

$$|i_1(z)|^2=\left(\frac{\eta V_1}{u_s}\right)^2\left(\frac{I_0}{u_s}\right)^2\left(\frac{\omega}{\omega_p}\right)^2\times\left\{g_1\sin^2\left[g_1^{\frac{1}{2}}\left(\frac{\omega_p z}{u_s}\right)\right]+\Delta(z)\right\}, \quad (28)$$

where

$$\Delta(z)\simeq\left(\frac{\omega}{\omega_p\sigma}\right)^2\frac{3}{2}\left\{-\frac{(2\omega_p z/u_s)^4}{4!}+\frac{2(2\omega_p z/u_s)^6}{6!}-\frac{3(2\omega_p z/u_s)^8}{8!}+\cdots\right\}. \quad (29)$$

The approximation made in determining $\Delta(z)$ given here is that of neglecting terms involving $1/\sigma^3$ and higher compared with $1/\sigma^2$ and neglecting terms in $(\omega/\omega_p\sigma)^4$ compared with $(\omega/\omega_p\sigma)^2$. Numerical checks indicate (29) is accurate within one or two percent for $\sigma\geqslant 1000$ and $\omega/\omega_p\leqslant 100$. Although the terms of Eq. (29) were carried only to the eighth power of $\omega_p z/u_s$, we might expect that the remaining terms will obey the indicated law of formation. If this is the case,

$$\Delta(z)\simeq\left(\frac{\omega}{\omega_p\sigma}\right)^2\frac{3}{2}\sum_{n=1}^{\infty}(-1)^n\frac{n(2\omega_p z/u_s)^{2n+2}}{(2n+2)!}. \quad (30)$$

Since

$$\sum_{n=1}^{\infty}(-1)^n\frac{n(2\omega_p z/u_s)^{2n+2}}{(2n+2)!}=\frac{1}{2}\left[\frac{\omega_p z}{u_s}\sin\frac{\omega_p z}{u_s}\cos\frac{\omega_p z}{u_s}-\sin^2\frac{\omega_p z}{u_s}\right], \quad (31)$$

$$\Delta(z)\simeq\left(\frac{\omega}{\omega_p\sigma}\right)^2\frac{3}{4}\left[\frac{\omega_p z}{u_s}\sin\frac{\omega_p z}{u_s}\cos\frac{\omega_p z}{u_s}-\sin^2\frac{\omega_p z}{u_s}\right]. \quad (32)$$

It is interesting to note that, to this order of approximation, $\Delta(z)$ goes through zero at the same points as the $\sin^2(\omega_p z/u_s)$. Thus the effect of the thermal spread

is to distort the amplitude of the standing wave of convection current. $\Delta(z)/(\omega/\omega_p\sigma)^2$ is plotted in Fig. 1.

DRIFTING STREAM INITIALLY POSSESSING FULL SHOT NOISE

To determine the effect of the thermal-velocity spread on noise, we consider a drifting stream which initially (at $z=0$) possesses full uncorrelated shot noise in each velocity class. To carry out this calculation properly, it is necessary to calculate the convection current produced by shot noise in each velocity class separately, find the square of its absolute magnitude, and then to sum over all the velocity classes to find the total mean-square noise current.

Again we assume a solution of the form of (19). Considering first the effect of the velocity class having the velocity u_i at $z=0$, the initial value of $F_1(z, u)$ is given by

$$F_1(0, u, u_i)=f_0(u, u_i) =[2eu_iF_0(u_i)\Delta f]^{\frac{1}{2}}[\delta(u-u_i)/u], \quad (33)$$

when Δf is the band width.

At this step in the calculation, we are allowing the ith velocity class at $z=0$ to possess full shot noise fluctuations in an amount which depends upon the amount of direct current carried by electrons of that velocity class. All of the other velocity classes at $z=0$ are considered to have no fluctuations. Since it has been assumed that the system is linear, we can consider the effect of each velocity class separately and then add up the mean-square noise current produced by the initial fluctuations in all the velocity classes by integrating the current produced by the ith velocity class over all values of u_i. The steps in this process are as follows:

$$F_1(z, u, u_i)=\sum_{n=0}^{\infty} f_n(u, u_i)z^n. \quad (34)$$

Then the convection current is

$$i_n(u_i, z)=\sum_{n=0}^{\infty} a_n(u_i)z^n, \quad (35)$$

where

$$a_n(u_i)=-\int_{-\infty}^{\infty} uf_n(u, u_i)du. \quad (36)$$

Multiplication of (35) by its complex conjugate then yields for the mean-square noise current density resulting from the fluctuations in the ith velocity class at $z=0$

$$\overline{i_n^2(u_i, z)}=\left\{\sum_{n=0}^{\infty} a_n(u_i)z^n\right\}\left\{\sum_{n=0}^{\infty} a_n^*(u_i)z^n\right\} =\sum_{n=0}^{\infty} b_{2n}(u_i)z^{2n}, \quad (37)$$

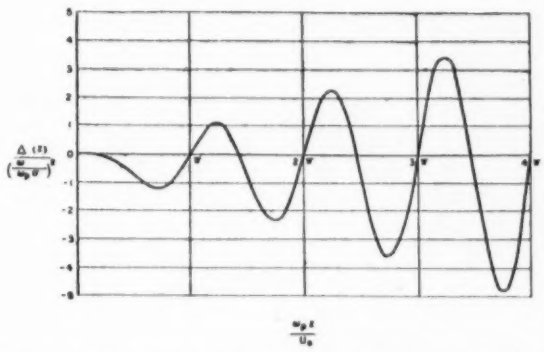


FIG. 1. A plot of $\Delta(z)/(\omega/\omega_p\sigma)^2$ given by Eq. (32). $\Delta(z)$ is the small correction factor for the ac current produced in an electron stream which has a velocity distribution.

where

$$b_{2n}=(-1)^n a_n^2+2\sum_{p=0}^{n-1}(-1)^p a_p a_{2n-p}. \quad (38)$$

The final step is the integration of (37) over u_i which yields for the total mean-square noise convection current density

$$\overline{i_n^2(z)}=2eI_0\Delta f\sum_{n=0}^{\infty} c_{2n}z^{2n}, \quad (39)$$

where

$$c_{2n}=\frac{1}{2eI_0\Delta f}\int_{-\infty}^{\infty} b_{2n}(u_i)du_i. \quad (40)$$

The first six coefficients of (39) are given in Appendix II. This series can be written in the form of a trigonometric function plus a small correction factor.

$$\overline{i_n^2(z)}=2eI_0\Delta f\left\{\cos^2\left[g_1^{\frac{1}{2}}\left(\frac{\omega_p z}{u_s}\right)\right]+\Delta_n(z)\right\}. \quad (41)$$

Neglecting terms in $1/\sigma^3$ and higher compared with $1/\sigma^2$, and neglecting terms in $(\omega/\omega_p\sigma)^4$ and higher compared with $(\omega/\omega_p\sigma)^2$,

$$\Delta_n(z)\simeq\left(\frac{\omega}{\omega_p\sigma}\right)^2\left\{\frac{1}{8}\left(\frac{\omega_p z}{u_s}\right)^4-\frac{33}{720}\left(\frac{\omega_p z}{u_s}\right)^6 +\frac{45}{8064}\left(\frac{\omega_p z}{u_s}\right)^8-\frac{427}{1{,}209{,}600}\left(\frac{\omega_p z}{u_s}\right)^{10}+\cdots\right\} \quad (42)$$

$\Delta_n(z)/(\omega/\omega_p\sigma)^2$ is plotted in Fig. 2 as a function of $\omega_p z/u_s$.

The principal point of interest is the value of $\Delta_n(z)$ at $\omega_p z/u_s=\pi/2$, since this is the minimum value of the mean-square noise convection current. Reference to Fig. 2 shows that at the first minimum

$$(\overline{i_n^2}/2eI_0\Delta f)|_{\min}=\Delta_n(\pi/2)\simeq(\omega/\omega_p\sigma)^2\cdot 0.25. \quad (43)$$

For oxide cathodes $KT_c/e\simeq 0.1$ volt. Then for $V_0=100$

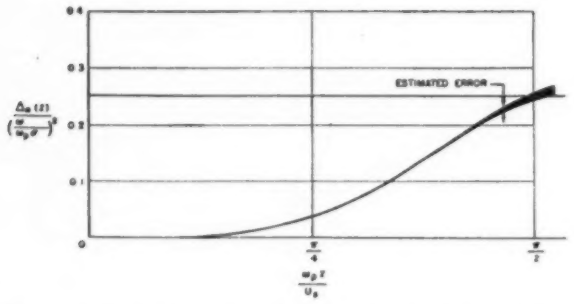


FIG. 2. A plot of $\Delta_n(z)/(\omega/\omega_p\sigma)^2$ given by Eq. (42). $\Delta_n(z)$ is the small correction factor to the mean-square noise current resulting from full shot noise at the input. The value of $\Delta_n(z)$ at $\omega_p z/u_s=\pi/2$ is the value of the first minimum of the noise current standing wave.

volts, a plasma frequency of 50 mc and a frequency of 5000 mc, the current at the first minimum is about 24 db below the maximum value.

The result of this calculation, showing that a standing wave of noise current is produced by pure uncorrelated shot noise, may be a little surprising to some of the workers in this field since many of us have believed that pure shot noise was the ultimate equilibrium state of an electron stream. Although a standing wave of noise current has been observed experimentally,[10] it was observed on an electron stream produced by space-charge limited rather than temperature-limited emission. Some of us believed that the action of the potential minimum was to produce a correlation among the noise fluctuations and that this correlation was the thing which permitted the production of a standing wave. The derivation given here demonstrates that such correlation is not essential and that the important factor in determining the behavior of the noise in the stream is space charge. It is interesting to note that the value of noise current at the first minimum is correctly predicted by considering a single-valued velocity stream for which one considers the standing wave produced by a current modulation equal to pure shot noise at the input plus the standing wave produced by velocity modulation equal to mean-square fluctuations in the average velocity of the stream associated with that pure shot noise.

CONCLUSIONS

A method of taking into account the thermal velocity spread in the formulation of electron beam problems has been described and two such problems have been solved using a power series. In the case of a drifting stream with velocity modulation, the effect of the thermal spread is to alter the amplitude and phase of the ac convection current density. In the case of an electron stream which initially contains full shot noise, it is found that a standing wave of convection current is produced but that the standing wave does not have perfect zeros. The value of the noise current at the first minimum increases as the square of frequency and with increasing velocity spread.

[10] C. C. Cutler and C. F. Quate, Phys. Rev. **80**, 875 (1950).

ACKNOWLEDGMENTS

The writer is indebted to J. R. Whinnery for many suggestions and for assistance in formulating and solving the problems, to A. V. Haeff for many stimulating discussions, and to W. L. Doyle for identifying the series of Eq. (31). The curves were calculated by Miss Margaret Gaine.

APPENDIX I

To calculate the initial value of the ac distribution function of an electron stream which has been given an ac velocity modulation we consider as a model an electron stream initially possessing the dc half-Maxwellian distribution of Eq. (12). In this model, the stream is allowed to pass through two parallel grids across which there appears the voltage $V_1e^{j\omega t}$. The grids are imagined to be perfectly permeable and so close that transit time effects cannot occur. The total distribution function including both the dc and ac parts to the left of the gap is given by

$$F(z,u,t)=F_0=2\alpha I_0 s(u-u_s)\exp[-\alpha(u^2-u_s^2)]. \quad \text{(A-1)}$$

To the right of the gap each velocity class will have acquired an additional amount of energy given by $eV_1e^{j\omega t}$. Thus, to the right of the gap

$$F(z,u,t)=2\alpha I_0 s[u-(u_s^2+2\eta V_1e^{j\omega t})^{\frac{1}{2}}] \times\exp[-\alpha(u^2-u_s^2-2\eta V_1e^{j\omega t})]. \quad \text{(A-2)}$$

To separate (2) into its dc and first-order ac parts, we expand the step function and the exponential in Taylor series as follows:

$$s[u-(u_s^2+2\eta V_1e^{j\omega t})^{\frac{1}{2}}]\simeq s\left(u-u_s-\frac{\eta V_1e^{j\omega t}}{u_s}\right)$$

$$\simeq s(u-u_s)-\frac{\eta V_1}{u_s}e^{j\omega t}\delta(u-u_s)+\cdots \quad \text{(A-3)}$$

$$\exp[-\alpha(u^2-u_s^2-2\eta V_1e^{j\omega t})] =\exp[-\alpha(u^2-u_s^2)](1+2\alpha\eta V_1e^{j\omega t}+\cdots). \quad \text{(A-4)}$$

Neglecting products of ac quantities we have for $F_1e^{j\omega t}$

$$F_1e^{j\omega t}=-2\alpha I_0\left\{\frac{\eta V_1e^{j\omega t}}{u_s}\delta(u-u_s) -2\alpha\eta V_1e^{j\omega t}s(u-u_s)\exp[-\alpha(u^2-u_s^2)]\right\}. \quad \text{(A-5)}$$

Comparison of (A-5) with (16) yields

$$F_1(0,u)=-(\eta V_1/u)(\partial F_0/\partial u) \quad (21)$$

since u_s can be replaced by u where it is multiplied by the delta-function.

APPENDIX II

Tabulation of the g_n's defined by Eq. (25) ($\sigma = eV_0/KT_c$):

$$g_1 \simeq 1 - \frac{3}{2\sigma} + \frac{15}{4\sigma^2} - \frac{105}{8\sigma^3} + \cdots,$$

$$g_2 \simeq \frac{1}{2\sigma} - \frac{9}{4\sigma^2} + \frac{87}{8\sigma^3} - \cdots,$$

$$g_3 \simeq \frac{1}{4\sigma^2} - \frac{9}{4\sigma^3} + \cdots,$$

$$g_4 \simeq \frac{1}{8\sigma^3} - \cdots,$$

etc.

Tabulation of the coefficients of Eq. (24):

$a_0 = 0,$

$a_1 = (\omega_p/u_s)g_1,$

$a_2 = j(\omega\omega_p/u_s^2)g_2,$

$$a_3 = -\frac{\omega^2\omega_p}{u_s^3}g_3 - \frac{1}{3!}\frac{\omega_p^3}{u_s^3}g_1^2,$$

$$a_4 = -j\frac{\omega^3\omega_p}{u_s^4}g_4 - j\frac{1}{3!}\frac{\omega\omega_p^3}{u_s^4}g_1g_2,$$

$$a_5 = \frac{\omega^4\omega_p}{u_s^5}g_5 + \frac{1}{5!}\frac{\omega^2\omega_p^3}{u_s^5}[2(3!)g_1g_3 + (2!)^2g_2^2] + \frac{1}{5!}\frac{\omega_p^5}{u_s^5}g_1^3,$$

$$a_6 = j\frac{\omega^5\omega_p}{u_s^6}g_6 + j\frac{1}{6!}\frac{\omega^3\omega_p^3}{u_s^6}[2(4!)g_1g_4 + 2(2!)3!g_2g_3] + j\frac{1}{5!}\frac{\omega\omega_p^5}{u_s^6}g_1^2g_2,$$

$$a_7 = -\frac{\omega^6\omega_p}{u_s^7}g_7 - \frac{1}{7!}\frac{\omega^4\omega_p^3}{u_s^7}[2(5!)g_1g_5 + 2(2!)4!g_2g_4 + (3!)^2g_3^2] - \frac{\omega^2\omega_p^5}{u_s^7}[2(3!)g_1^2g_3 + 3(2!)^2g_1g_2^2 + 3!g_1^2g_3] - \frac{1}{7!}\frac{\omega_p^7}{u_s^7}g_1^4,$$

$$a_8 = -j\frac{\omega^7\omega_p}{u_s^8} - j\frac{1}{8!}\frac{\omega^5\omega_p^3}{u_s^8}[2(6!)g_1g_6 + 2(2!)5!g_2g_5 + 2(3!)4!g_3g_4] - j\frac{1}{8!}\frac{\omega^3\omega_p^5}{u_s^8}[4(2!)3!g_1g_2g_3 + (2!)^3g_2^3 + 3(4!)g_1^2g_4] - j\frac{1}{7!}\frac{\omega\omega_p^7}{u_s^8}g_1^3g_2.$$

Coefficients of Eq. (39):

$c_0 = 1,$

$c_2 = -(\omega_p^2/u_s^2)g_1,$

$$c_4 = \frac{8}{4!}\frac{\omega_p^4}{u_s^4}g_1^2 + \frac{2(3!)}{4!}\frac{\omega^2\omega_p^2}{u_s^4}g_3,$$

$$c_6 = -\frac{32}{6!}\frac{\omega_p^6}{u_s^6}g_1^3 - \frac{1}{6!}\frac{\omega^2\omega_p^4}{u_s^6}[204g_1g_3 - 72g_2^2] - \frac{2(5!)}{6!}\frac{\omega^4\omega_p^2}{u_s^6}g_5,$$

$$c_8 = \frac{128}{8!}\frac{\omega_p^8}{u_s^8}g_1^4 + \frac{1}{8!}\frac{\omega^2\omega_p^6}{u_s^8}[1548g_1^2g_3 - 648g_1g_2^2] + \frac{1}{8!}\frac{\omega^4\omega_p^4}{u_s^8}[2592g_3^2 + 7200g_1g_5 - 5184g_2g_4] + \frac{2(7!)}{8!}\frac{\omega^6\omega_p^2}{u_s^8}g_7,$$

$$c_{10} = -\frac{512}{10!}\frac{\omega_p^{10}}{u_s^{10}}g_1^5 - \frac{1}{10!}\frac{\omega^2\omega_p^8}{u_s^{10}}[9228g_1^3g_3 - 4104g_1^2g_2^2] - \frac{1}{10!}\frac{\omega^4\omega_p^6}{u_s^{10}}[72{,}720g_1^2g_5 + 33{,}696g_1g_3^2 - 62{,}208g_1g_2g_4 - 1296g_2^2g_3] - \frac{1}{10!}\frac{\omega^6\omega_p^4}{u_s^{10}}[473{,}760g_1g_7 - 339{,}840g_2g_6 - 144{,}000g_4^2 - 305{,}280g_3g_5] - \frac{2(9!)}{10!}\frac{\omega^8\omega_p^2}{u_s^{10}}g_9.$$

Magnetically Induced Ultrasonic Velocity Changes in Polycrystalline Nickel*

S. J. JOHNSON,† *Andersen Laboratories, Incorporated, West Hartford, Connecticut*

AND

T. F. ROGERS,‡ *Air Force Cambridge Research Center, Cambridge, Massachusetts*

(Received January 30, 1952)

Measurements of magnetically induced velocity changes in polycrystalline nickel rods are discussed. Shear and compressional ultrasonic propagation modes are used throughout the frequency range of 1–10 megacycles. The polycrystalline measurements are compared to recently published single nickel crystal values. The ΔE effect is found to decrease with increasing frequency throughout the megacycle region, showing general agreement with low frequency measurements.

THIS paper is presented to report upon our continuing experiments concerning certain magnetoacoustic effects in polycrystalline nickel, the initial results of which were reported[1] one year ago. At that time, as part of a broader program of study of ultrasonic propagation in solid media, changes in both the ultrasonic velocity and absorption induced in pure nickel rods by the application of a magnetic field to this transmission medium were noted; a maximum change of some 3 percent was induced in the velocity of propagation of shear 1-megacycle ultrasonic waves by placing the rod in a longitudinal field of up to some 1500 oersteds with a profound decrease in absorption being noted as taking place simultaneously. Measurements of these phenomena have been continued, and the results of our velocity change studies are reported here; similar changes found in other ferromagnetic materials as well as associated absorption changes will be presented at a later date.

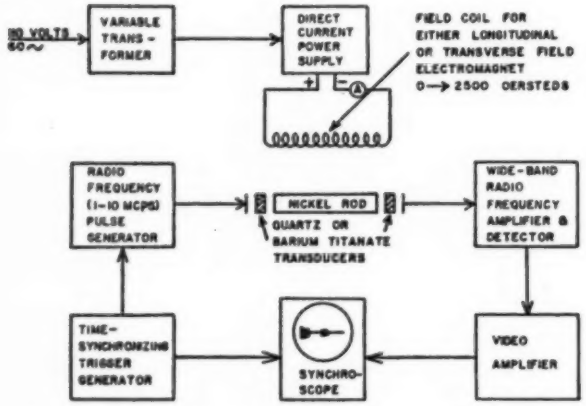

110 VOLTS 60~
VARIABLE TRANS-FORMER
DIRECT CURRENT POWER SUPPLY
FIELD COIL FOR EITHER LONGITUDINAL OR TRANSVERSE FIELD ELECTROMAGNET 0→2500 OERSTEDS
RADIO FREQUENCY (1–10 MCPS) PULSE GENERATOR
NICKEL ROD
QUARTZ OR BARIUM TITANATE TRANSDUCERS
WIDE-BAND RADIO FREQUENCY AMPLIFIER & DETECTOR
TIME-SYNCHRONIZING TRIGGER GENERATOR
SYNCHRO-SCOPE
VIDEO AMPLIFIER

FIG. 1. Functional block diagram of the method for measuring magnetically induced velocity changes in a polycrystalline nickel rod.

Since our initial reports our measurement techniques have been modified and extended in several ways:

(a) Both shear (transverse) and compressional (longitudinal) modes of propagation have been employed.

(b) Fields parallel to the rod axis—here called longitudinal fields—have continued to be employed. In addition we have also used transverse fields applied normal to the rod axis with various orientations relative to the direction of ultrasonic particle motion.

(c) The frequency range has been extended to cover the 1- to 10-megacycle region.

The measurement method (Fig. 1) consists essentially in coupling a short (about 10 microseconds in most cases) pulse of acoustic energy into the nickel rod by means of either a quartz or barium titanate ($BaTiO_3$) transmitting transducer (used for shear and compressional modes, respectively) driven by a variable center frequency electronic pulse generator. The received delayed sonic pulse packet is reconverted by means of a similar receiving transducer into an electrical signal and is amplified in either a wide band video or I.F. type amplifier. The delayed electrical signal is displayed upon a synchroscope whose sweep is phase synchronized to the starting time of the pulse generator, with variations in pulse delay time resulting from the electromagnet field application being measured directly upon the synchroscope face using an accurately calibrated electronic strobe.

Larger rod diameters, varying from some 1.25 to 4 centimeters, have been used to reduce interference from rod wall reflection and propagation mode conversion effects[2] by means of achieving a higher diameter-to-wavelength ratio. Even under high D/λ conditions, however, accurate measurements upon polycrystalline nickel are still made difficult, especially at the higher frequencies where the effects to be measured are small, by the elastic anisotropy of the individual nickel crystals making up the sample; the more or less random array of these anisotropic crystals in the polycrystalline rod gives rise to a velocity difference (the anisotropy

* Supported by the Air Force Cambridge Research Center. Presented at the Chicago meeting of the American Physical Society, October, 1951.

† Formerly at Crystal Research Laboratories, Inc., Hartford, Connecticut.

‡ Now at Massachusetts Institute of Technology.

[1] T. F. Rogers and S. J. Johnson, J. Appl. Phys. **21**, 1067 (1950). T. F. Rogers and S. J. Johnson, Proc. Inst. Radio Engrs. **39**, 307A (1951).

[2] Hughes, Pondrom, and Mims, Phys. Rev. **75**, 1552 (1950). W. P. Mason, *Peizoelectric Crystals and Their Application to Ultrasonics* (D. Van Nostrand, Company, Inc., New York, 1950), p. 430.

factor, $2C_{44}/(C_{11}-C_{22})$, is 2.63 in nickel) for various directions of particle motion and directions of propagation which results in velocity dispersion and interference effects in the received pulse components due to scattering by the individual grains. This difficulty has been reduced by making measurements upon the individual cycles of the received wave fronts rather than the envelope of the detected pulse.

Our initial study showed that the observed changes in delay, because of their magnitude relative to any magnetostrictive or thermal changes, can only be caused by a marked change in the ultrasonic velocity of propagation within the rod and indicated that this velocity change was to be attributed to a variation of Young's modulus E under the applied field stress, since it will be recalled that the sonic velocity is proportional to $E^{\frac{1}{2}}$. From static and low frequency measurements of the free longitudinal vibration of ferromagnetic bars in the

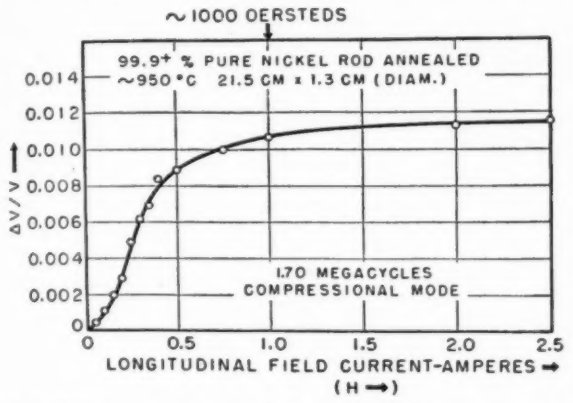


FIG. 2. Change in ultrasonic velocity in polycrystalline nickel with an increasing applied magnetic field. One ampere of electromagnet field current is equivalent to a field strength of about 1000 oersteds.

kilocycle region made previously,[3] this change in Young's modulus with magnetization has come to be called the "ΔE effect." Our further work confirms this interpretation of our experimental results, as do the papers reporting upon ultrasonic measurements of nickel by Mason and his co-workers[4–6] at the Bell Telephone Laboratories and of mild steel by Petralia[7] at the University of Bologna, all of which have appeared since that time.§

It is to be noted that a nickel crystal is both magnetically and elastically anisotropic, and this forms the basis for the effects noted here. That nickel is magnetically anisotropic has been known for some time from measurements upon single crystals, and measurements of the ultrasonic velocity along different crystalline axes[8] has shown a definite difference in the various velocity values. Upon being subjected to a magnetic field, it is found that the magnetically saturated values of the velocity (from which the single crystal elastic constants, the average elastic constants for a polycrystalline structure, and the effective value of E may be deduced) differ from those obtaining in the unmagnetized case. They also differ slightly, for any given propagation and particle motion direction, with the direction of the applied field, but this "morphic" effect[4] is overshadowed in the polycrystalline case, only the average value being of importance in our measurements. Table I summarizes such measurements taken upon a single nickel crystal; the direction of propagation in all cases is along the [110] direction and the frequency is 10 megacycles.

TABLE I. Summary of the single nickel crystal magnetically saturated acoustic velocities.[a]

Transmission mode	Particle motion	Field direction	V_s/V	$Av\Delta V_s/V$
Shear	[110]	[110]	1.0224	0.0242
Shear	[110]	[001]	1.0260	
Shear	[001]	[1$\bar{1}$0]	1.0130	0.0135
Shear	[001]	[001]	1.0140	
Longitudinal	[110]	[1$\bar{1}$0]	1.0061	0.0062
Longitudinal	[110]	[001]	1.0063	

[a] After Bozorth, Mason, and McSkinnin, reference 6.

Using either a transverse or a longitudinal magnetic field, a pronounced change in ultrasonic velocity is found for both shear and compressional waves, with low frequency (1-megacycle) saturation values, $\Delta V_s/V$ approximating 0.5–3 percent, depending upon the specimens used and their previous heat treatment. The change in velocity with applied field, an increase in all cases and independent of field direction, is large for small applied field (Fig. 2) i.e., for $H \leqq 200$–300 oersteds. With field increases above this region the increase in $\Delta V/V$ is less rapid, and for $H \cong 1500$ oersteds saturation is approached and $\Delta V/V$ becomes independent of field strength. Upon reducing the field to zero, the original velocity does not usually obtain, but further application of a demagnetizing (ac) field produces essentially the original velocity, V. Figure 3 compares a representative polycrystalline $\Delta V/V$ characteristic with an analogous measurement for the single nickel crystal case. Except for the absolute values of the change, there is a similarity in the results, as might well be expected.

Of note, however, is the fact that in many cases the value of $\Delta V/V$ which obtains in the limit of zero field as the field is reduced from saturation is appreciably

[3] R. M. Bozorth, *Ferromagnetism* (D. Van Nostrand Company, Inc., New York, 1951), p. 684.
[4] W. P. Mason, Phys. Rev. **82**, 715 (1951).
[5] W. P. Mason, Phys. Rev. **83**, 683 (1951).
[6] Bozorth, Mason, and McSkimin, Bell System Tech. J. XXX, Pt. I, 970 (1951). (Dr. Mason very kindly provided us with a prepublication copy of this paper.)
[7] S. Petralia, Atti acca. nazl. Lincei (R. C. Cl. Sci. Fis. Mat. Nat.) **10**, 146 (1951).
§ *Note added in proof:*—Since submitting this paper the recent work of J. de Klerk [Nature **168**, 963 (1951)] has reached our attention; both changes in velocity and attenuation have been measured in ferromagnetic rods, confirming our earlier observations.
[8] Bozorth, Mason, McSkimin, and Walker, Phys. Rev. **75**, 1954 (1949).

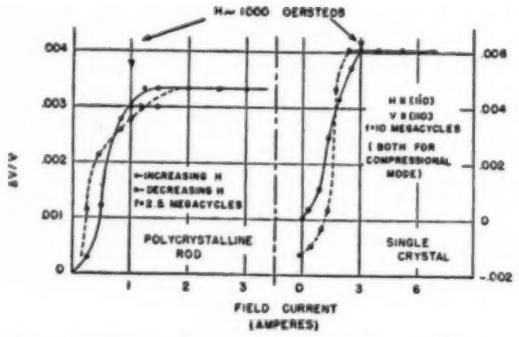


FIG. 3. Comparison of a representative polycrystalline measurement with an analogous measurement for a single nickel crystal; single crystal values are taken from Fig. 9 of the paper by Mason (reference 4).

larger than that measured under initial zero field conditions. The polycrystalline measurements by Bozorth, *et al.*,[9] which also show this same departure to a higher $\Delta V/V$ value for the reduced field, are in contrast to the single crystal measurements[5] wherein the reduced field values in all cases are lower than the initial zero field ones. While such lower $\Delta V/V$ values for the final zero field may be explained upon the basis that the domain free energy is lower for other than the demagnetized state, a higher final value is not so easily understandable.

Application of a saturation field over many cycles has been noted upon occasion to produce a permanent change in the ΔV properties, a decrease in $\Delta V_s/V$ being most generally the case. This "aging" of the metal is also found by Petralia[10] in his work with steel in the 0.8–1.7 megacycles region.

The $\Delta V_s/V$ values reached in the case of shear waves in a longitudinal field are generally 2–3 times that found for compressional waves (Fig. 4); in the measurements of the saturation single crystal $\Delta V_s/V$ values the two shear wave values were about 2 and 4 times, respectively, those measured for the compressional waves with the same conditions of propagation axis and field direction (see Table I).

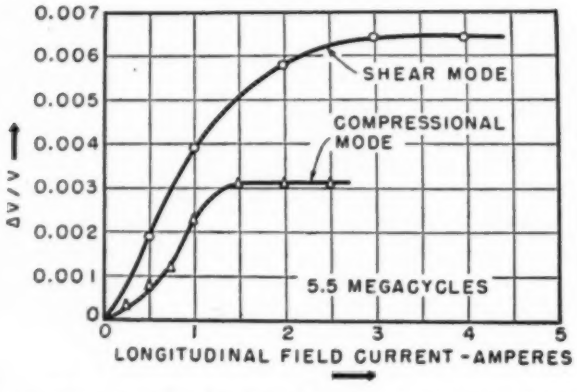


FIG. 4. Comparison of the change in ultrasonic velocity with applied field for shear and compressional propagation modes.

In attempts to obtain higher values of ΔE, certain of the rods were annealed at high temperatures (up to about 900°C in some cases) since static measurements upon nickel by Köster[11] indicate that such elevated temperature annealing produces an increase in the ΔE effect of a factor of five or more. Increases in the saturation values produced by this treatment have been obtained, but, unfortunately, the heat treatment results in such an inordinately high increase in ultrasonic absorption within the rod (which other evidence indicates is strongly dependent upon crystalline size and orientation in polycrystalline media in the megacycle region) that received signals could only by observed under high field conditions, thus making other than saturation $\Delta V/V$ measurements impossible. This increase in absorption agrees with Köster's static measurements[12] in which the absorption increased proportional to the square (approximately) of the ΔE effect increase produced by such heat treating.

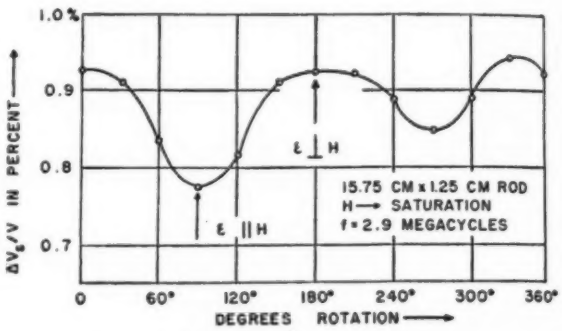


FIG. 5. The variation of shear mode propagation saturation $\Delta V/V$ values in a transverse field as the rod, and consequently, the ultrasonic particle motion ϵ is rotated relative to the field direction.

With the use of shear waves and a transverse field, a variation in $\Delta V_s/V$ values was noted which is related to changes in the orientation of the field with respect to the ultrasonic particle motion in the rod (Fig. 5). The two shear wave single crystal values of $\Delta V_s/V$ taken for different particle motion with respect to the field (Table I) differ by a factor of two, and any other but random orientation would therefore allow such variations to be noted in the polycrystalline rod. The measurements, therefore, may be assumed to indicate that the crystalline structure is not arrayed completely at random within the rod, but that rather the working of the metal prior to our use has resulted in an elastic anisotropy, i.e., a preferred orientation of the crystalline axis direction; the observed changes are analogous to those observed in the magnetic properties of anisotropic sheets of ferromagnetic materials for

[9] Reference 6, Fig. 10.
[10] Reference 7, p. 148.
[11] W. Köster, Z. Metallkunde 35, 57 (1943).
[12] W. Köster, Z. Metallkunde, 35, 246 (1943).

various angles relative to the fabrication rolling direction.[13]

Static and low frequency (the kilocycle region) measurements of the ΔE effect in nickel by other workers[14] showed that changes of as much as about 20 percent may be found at room temperature. Our initial measurements at one megacycle indicated a maximum saturation ΔE effect of only some 6.4 percent, with many specimens measuring considerably less, and, as in the case of the Bell Telephone group (who deduced maximum 10-megacycle ΔE changes of only some 3 percent from single crystal measurements)[6] a search for a dependence of this effect upon frequency was indicated and independently conducted in the megacycle region. A pronounced dependence has been found in all cases of field direction and mode of propagation. The change in velocity and, consequently, the ΔE effect has been found to decrease with increasing frequency throughout the entire 1- to 10-megacycle region, thereby extending and, in general, corroborating recently reported[5,6] low frequency (10–100 kilocycle) measurements on polycrystalline nickel.

Mason[5] has derived an equation for the frequency dependence of the ΔE effect for a single domain of the form

$$\frac{\Delta E}{E}=\frac{\Delta E}{E_0}\left(\frac{1}{1+f^2/f_0^2}\right), \tag{1}$$

where $\Delta E/E_0$ is the static (low frequency) value for the maximum change of E relative to the unmagnetized E_0 value, and f_0 is the relaxation frequency (dependent upon domain size and other parameters) which, for the rods used in his experiments, approximated 150 kilocycles. At frequencies such that $f \gg f_0$,

$$\frac{\Delta E}{E}=\left(\frac{\Delta E}{E_0}\right)\Big/\left(\frac{f}{f_0}\right)^2. \tag{2}$$

With a given value of f_0 (in the region of $1-2\times10^5$) and, since the acoustic velocity is proportional to $E^{\frac{1}{2}}$, the velocity relationships in the 1–10 megacycle range should be of the form

$$\Delta V/V=K/f. \tag{3}$$

From our measurements of both shear and compressional mode saturation velocity changes (Fig. 6), it is apparent that the frequency dependence predicted by the simple single domain size model does not obtain.

[13] See reference 3, pp. 12, 42, and 587; also R. F. S. Hearmon, Revs. Modern Phys. 18, 409 (1946).

[14] Reference 3, Chapter 13, especially p. 697.

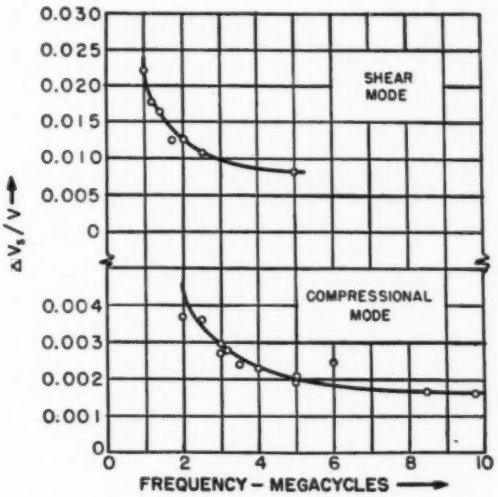


FIG. 6. Ultrasonic frequency dependence of saturated $\Delta V/V$ values in the 1–10 megacycle region for a longitudinal magnetic field. The upper curve is for shear mode propagation, the lower for the compressional mode.

While at these frequencies the velocity changes are appreciably smaller than in the kilocycle region and, as a consequence are more difficult to measure with absolute accuracy (the velocity changes representing time delay changes of a few tenths of a microsecond), still, the relative march of values is significant in this respect; certainly, Eq. (1) does not admit of the effect falling off less rapidly with increasing frequency for any range of frequencies. Taking into consideration the fact that our polycrystalline sample contains a distribution of domain sizes, however, the measured values are not too surprising, and in fact they bear a marked similarity to the kilocycle polycrystalline measurements.[6]

It is to be noted that the low $\Delta E/E$ values in the case of the two relatively inactive single crystals and their equivalent $\Delta V/V$ values, and our polycrystalline measurements lie close to the limits of the measurement accuracy (0.1 percent) claimed in a recent paper by Levy and Truell,[15] and it is conceivable that a small increase in average domain size in the crystal used in their measurements relative to the ones noted here might explain their failure to observe any change in velocity at 5 megacycles and higher. The $\Delta E/E$ values obtained both for single and polycrystalline cases are of course greatly dependent upon the history of the material under test with domain size and, possibly, internal strain effects being dominant.

The authors wish to thank W. A. Andersen, A. Ganio, and V. Miner for their help in this work.

[15] S. Levy and R. Truell, Phys. Rev. 83, 668 (1951).

Heat Conduction in Alloys at Low Temperatures*

I. ESTERMANN† AND J. E. ZIMMERMAN‡
Carnegie Institute of Technology, Pittsburgh, Pennsylvania
(Received February 18, 1952)

With a view to studying the mechanism of heat conduction in low conductivity alloys, a method has been devised by which the thermal conductivity of relatively small samples ($\frac{1}{8}$- to $\frac{1}{4}$-inch diameter, 1 to 2 inches long) of various materials can be measured in the temperature regions obtainable with liquid nitrogen, liquid hydrogen, and liquid helium.

Preliminary measurements on several commercial alloys (monel, inconel, and stainless steel) gave Wiedemann-Franz ratios several times greater than the theoretical value of 2.45×10^{-8} watt-ohm/deg^2, the deviation being greater for annealed than for cold-worked specimens. This has been interpreted in terms of appreciable lattice conduction of heat in these alloys.

Following these results, samples of an alloy of 90 percent copper 10 percent nickel were prepared with varying amounts of cold-work and with different grain sizes. Results with these samples were similar to those obtained with monel and inconel and confirm the hypothesis of lattice conduction. They also give a qualitative indication of the effectiveness of cold-work in limiting the lattice conduction of heat.

INTRODUCTION

IT has been known for a long time, being first observed by Wiedemann and Franz in 1853, that for pure metals with high electrical conductivity the ratio of thermal to electrical conductivity is approximately constant, and it was further observed by Lorenz in 1881 that this constant is about proportional to the absolute temperature. This so-called Wiedemann-Franz law has since been derived theoretically and has been found to be valid within a few percent in the region of room temperature and above for a large number of metals and high conductivity alloys.[1]

In recent years extension of the measurements to low temperatures have revealed large deviations from this simple law. For the pure metals referred to above it has been found that the Wiedemann-Franz ratio $K/\sigma T$, where K is the thermal conductivity, σ the electrical conductivity, and T the absolute temperature, usually becomes considerably less than the theoretical value at temperatures of the order of one-fourth the Debye characteristic temperature θ. This behavior has been predicted qualitatively by Wilson[2] and others,[3,4] who have shown in fact that for perfectly pure metals with no lattice defects $K/\sigma T$ should be proportional to T^2 at low temperatures.

Alloys, semiconductors, and pure metals having relatively low electrical conductivity usually show quite a different type of deviation from the simple law, $K/\sigma T$ for these materials being generally greater than the theoretical value at temperatures of the order of $\theta/4$ or less. Bismuth and antimony, which fall into the last-mentioned class, have been subjected to measurements down to liquid-hydrogen temperatures,[5] and bismuth has also been considered theoretically by Makinson.[3] It has been suggested on theoretical grounds and confirmed experimentally that in these metals lattice conduction of heat is appreciable compared to conduction by free electrons; in particular it has been shown experimentally that $K/\sigma T$ depends markedly upon grain size and also, in the case of bismuth, on magnetic field. Both of these results are taken as an indication of the presence of lattice conduction.

In the case of metallic alloys, and in particular disordered alloys, since a perfectly ordered alloy has the properties of an ideal metal, it is also reasonable to expect that lattice and electronic conduction might be of the same order of magnitude, especially at rather low temperatures where lattice conduction has a maximum, theoretically, and electronic conduction is small as compared to that in pure metals. The importance of lattice conduction in alloys has been demonstrated in only a few instances, however, and experimental research should be useful for this and other reasons. Part of the work described in this paper was undertaken for the very practical purpose of providing data on commercial alloys useful in the design of low temperature equipment. Some of the data have already been put to use.[6] Another part of more fundamental significance is a systematic investigation of the effect of cold-work on thermal and electrical conductivities of a particular alloy and the theoretical implications. Such an investigation is especially interesting because in many disordered alloys electronic conduction is limited almost entirely by the disorder and is not appreciably affected by such long-range parameters as grain size and lattice dislocations; hence, one can study the effect of grain size and lattice dislocations with the electronic conduction, in effect, held constant. Finally, a study has been made of the temperature variation of thermal

* Assisted by the ONR. This is an abstract of the thesis submitted by J. E. Zimmerman in partial fulfilment of the requirements of the degree of Doctor of Science.
† At present with the ONR, Washington, D. C.
‡ Gerard Swope Fellow at the time this work was started.
[1] H. Landolt and R. Börnstein, *Physikalisch-Chemische Tabellen* (Verlag Julius Springer, Berlin, 1931), Vol. II, p. 1023; (1936), Vol. III, p. 1998.
[2] E. H. Wilson, Proc. Cambridge Phil. Soc. **33**, 371 (1937).
[3] R. E. B. Makinson, Proc. Cambridge Phil. Soc. **34**, 474 (1938).
[4] E. H. Sondheimer, Proc. Roy. Soc. (London) **203**, 75 (1950).
[5] A. Eucken and O. Neumann, Z. Phys. Chem. **111**, 431 (1924).
[6] A. Wexler, J. Appl. Phys. **22**, 1463 (1951).

conductivity in the neighborhood of absolute zero, where theory predicts the vanishing of lattice conduction and the validity of the Wiedemann-Franz law.

THEORY

Theoretical analysis of heat conduction in alloys usually takes into account two conducting mechanisms operating in parallel and mutually interfering with each other, (1) the conduction by free or conduction electrons and (2) the conduction by the crystal lattice. The thermal conductivity K may then be written as the sum of two components, the electronic conductivity K_e and the lattice conductivity K_l, or

$$K=K_e+K_l. \tag{1}$$

In the approximation where the conduction electrons are considered to be perfectly free and are scattered chiefly by impurities and lattice defects but not by thermal vibrations, K_e is given by

$$K_e=\sigma_0 L_0 T=1/3(\pi k/e)^2\sigma_0 T, \tag{2}$$

where σ_0 is the residual electrical conductivity, T is the absolute temperature, k is Boltzmann's constant, e is the electronic charge, and $L_0=1/3(\pi k/e)^2=2.45\times10^{-8}$ watt/ohm/deg^2. This is the well-known law of Wiedemann and Franz, the constant L_0 being independent of such parameters as magnetic field, grain size, and degree of cold-work. At temperatures above the Debye characteristic temperature of the material in question this law is valid for thermal scattering as well as for impurity and defect scattering and may be written $K_e=\sigma L_0 T$, where σ is the electrical conductivity at the temperature T. The requirement that the free electrons be scattered chiefly by impurities and lattice defects is equivalent to requiring that σ be essentially equal to σ_0.

A theoretical form for the lattice conductivity K_l has been derived by Makinson, namely

$$K_l=3Nvk\left(\frac{T}{\theta}\right)^3\int_0^{\theta/T}\frac{z^4 l(Tz,T)dz}{(e^z-1)(1-e^{-z})};\quad z=h\nu/kT, \tag{3}$$

where N is the number of atoms per unit volume, v is the velocity of sound, θ is the Debye characteristic temperature of the material, ν is the frequency of lattice vibration, and k and h are Boltzmann's and Planck's constants, respectively. The function $l(Tz, T)=l(\nu, T)$ will be called the lattice wave mean free path. Its magnitude and temperature and frequency dependence are determined by the mechanisms responsible for scattering the acoustic waves in the lattice. A discussion of scattering by grain boundaries, by free electrons, by impurities, and by mutual intereference is given in the original article. The role of dislocations and other lattice defects induced by cold-working has not been considered.

The equation for K_l may be understood qualitatively in terms of the kinetic theory of gases and the specific heat theory of Debye. The fundamental formula for thermal conductivity of a gas is[7]

$$K=(1/3)vlC_v, \tag{4}$$

where C_v is the specific heat per unit volume at constant volume. Substituting for C_v the integrand in Debye's equation for the specific heat and integrating, Makinson's equation results.

It is not possible to calculate the numerical value of K_l as a function of temperature, since the mean free path l of the lattice waves has not been calculated precisely. Makinson's work shows, however, that in the neighborhood of absolute zero l is a constant determined by grain size, or in some cases by the dimensions of the specimen itself, so that K_l is proportional to C_v, that is, to T^3, in this temperature region. K_l does not increase indefinitely with temperature, but rather approaches a maximum and then decreases again at higher temperatures as a result of mutual interference of the lattice waves at large amplitudes. Its value in the intermediate temperature region is determined more or less by all of the previously mentioned scattering mechanisms.

If we write $K/\sigma T=L'$, then from Eqs. (1) and (2)

$$L'/L_0=K/\sigma L_0 T=\sigma_0/\sigma+K_l/\sigma L_0 T,$$

and with the approximation $\sigma_0=\sigma$ this becomes

$$L'/L_0=1+K_l/\sigma_0 L_0 T=1+K_l/K_e. \tag{5}$$

Thus if K_l is small or absent, L' is equal to L_0 and the Wiedemann-Franz law is valid; in general, however, L' is greater than L_0.

EQUIPMENT AND EXPERIMENTAL ARRANGEMENT

The problem of direct measurement of thermal conductivity at low temperature is in principle a rather simple one. As shown in Fig. 1, one end of a cylindrical specimen is attached to a cryostat (constant-temperature heat reservoir), the other end is attached to a heater, and the temperatures of two intermediate points

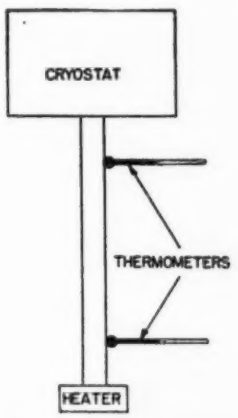


FIG. 1. Basic set-up.

[7] L. Loeb, *Kinetic Theory of Gases* (McGraw-Hill Book Company, Inc., New York, 1927), p. 240.

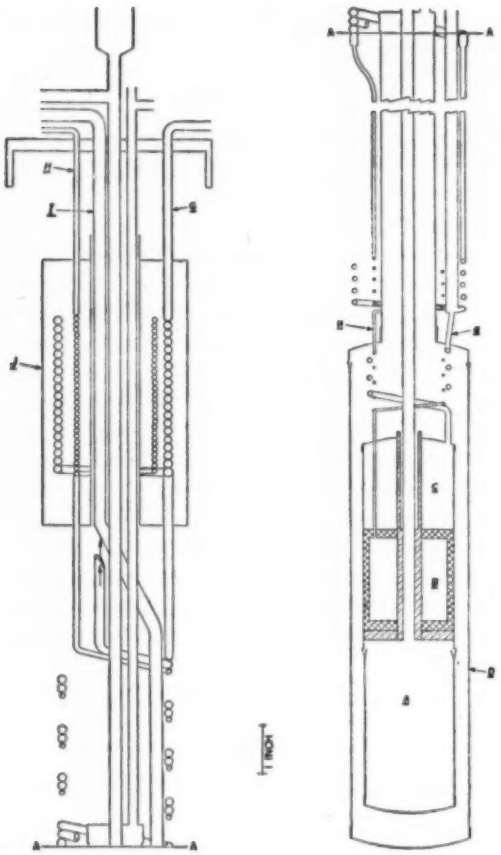


FIG. 2. Cryostat schematic.

are measured. The average thermal conductivity is then given by

$$K = WD/A(T_2 - T_1), \tag{6}$$

where W is the heater power, D the distance between thermometers, A the cross-sectional area of the specimen, and T_1 and T_2 the thermometer readings. At temperatures below 100°K, one need not be too concerned with errors due to heat radiation nor, if measurements are carried out in a vacuum, to gaseous conduction.

Equipment capable of achieving and measuring the required low temperatures is fairly complex, however. The rest of this section contains a description of the cryogenic and thermometric equipment built for operation between 3 and 80°K, or, more precisely, for operation at temperatures obtainable with liquid nitrogen, liquid hydrogen, and liquid helium. A final paragraph is devoted to the heater and heater circuit.

The Cryostat

The cryostat, a cross-sectional view of which is given in Fig. 2, produced its own liquid helium by the Simon single-expansion process, and functioned also as a cryostat at liquid nitrogen and liquid hydrogen temperatures with the aid of external facilities for liquefying the nitrogen and hydrogen. The principle and practice of liquefying helium by the single-expansion method has been dealt with extensively by Simon himself[8] and others, so this discussion will be confined to the mechanical design, the particular method of operation, and the performance data.

A complete "flow sheet" showing connections between the cryostat and the auxiliary equipment is given in Fig. 3. Corresponding parts in Figs. 2 and 3 have been lettered for reference. The working space in which the test specimen was mounted is designated by A.

For operation at hydrogen and helium temperatures, the metal Dewar flask E was filled with liquid nitrogen and the glass Dewar F, after pre-cooling with liquid nitrogen, was filled to the indicated level with liquid hydrogen. Hydrogen gas at a pressure of a few pounds per square inch above atmospheric was then admitted through the tube G to container C, where it condensed through the action of a heat transfer gas (also hydrogen) in the vacuum can D. At the same time helium at 120 to 150 atmospheres pressure was admitted to the helium bomb B through tube H. After container C was filled with liquid hydrogen (about half an hour after condensation began), the transfer gas was pumped out and the apparatus was ready for a series of measurements at liquid hydrogen temperatures. For producing liquid helium, the vapor pressure of the hydrogen in C was lowered adiabatically by pumping through tube I. After about an hour of pumping, a limiting vapor pressure of about 2 mm, corresponding to a temperature of around 10°K, was attained. At this point, the helium gas in B was expanded, over a period of about ten minutes, to atmospheric pressure. 30 to 40 cm³ of liquid helium was produced by this expansion, and the vaporation rate, 0.08 to 0.10 cm³ of liquid per minute, was such that measurements at liquid helium temperatures could be carried on for a period of several hours.

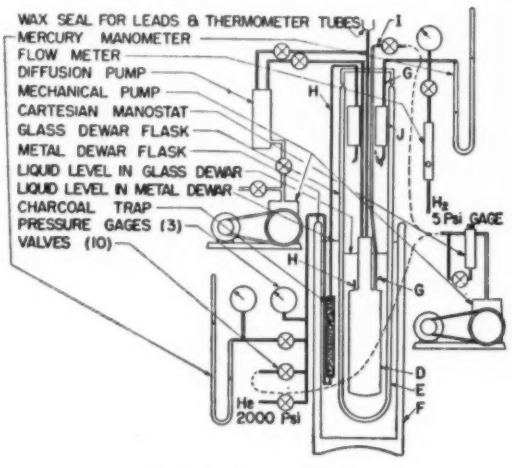


FIG. 3. Flow diagram.

[8] F. E. Simon and G. L. Pickard, Proc. Phys. Soc. (London) 60, 405 (1948).

Measurements at hydrogen temperatures were usually made, as stated above, before expansion of the helium. Sometimes, however, hydrogen temperature measurements were made after the liquid helium had all evaporated. About two liters of liquid hydrogen were required to fill the glass Dewar to the indicated level, and the rate of evaporation, about 0.8 cm³ of liquid per minute, was long enough to permit a complete set of measurements to be made with time out for meals, etc.

The pre-cooler can J was intended to be filled with liquid nitrogen for the purposes of pre-cooling the incoming hydrogen and helium gases, and of reducing the rate of evaporation of hydrogen in the glass Dewar. No appreciable saving in liquid hydrogen resulted from its use, however, and after the first two or three runs the pre-cooler was left unfilled.

Nitrogen temperatures were achieved in the working space of the cryostat by filling the glass Dewar with liquid nitrogen and then condensing nitrogen in container C, again using hydrogen as the heat transfer medium. The helium bomb was not used. After pumping out the transfer gas, the temperature could be varied between 58 and 80°K by varying the vapor pressure of the nitrogen in C.

In general, the vapor pressure of the cryostat liquid (helium, hydrogen, or nitrogen) was taken as a measure of the temperature. At first, the vapor pressure was controlled by manually adjusting the pumping rate. This was difficult to do with precision, and later measurements were made with a Cartesian manostat§ in the pumping line to control the vapor pressure. Vapor pressures were measured by means of mercury manometers connected as shown in Fig. 3.

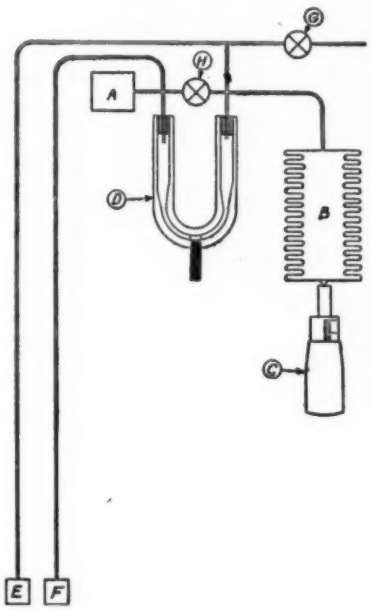


FIG. 4. Differential thermometer.

§ Emil Greiner Company model No. 6.

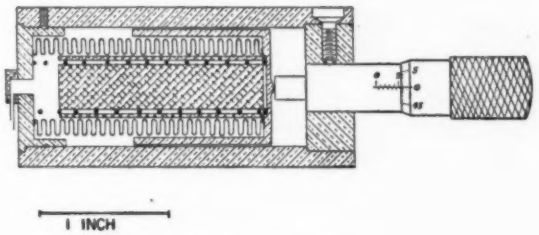


FIG. 5. Bellows assembly.

The Thermometer

For use in measuring thermal conductivities of small samples of various materials at temperatures ranging from the liquid nitrogen down to the liquid helium region, a new type of differential perfect-gas thermometer capable of detecting temperature differences of the order of 0.01 percent of the absolute temperature was designed and constructed.

A schematic diagram of the thermometer is given in Fig. 4. E and F are the gas thermometer bulbs, having equal volumes v of about 1/2 cm³, which in this application are thermally attached to two points on the sample. Each of these is connected to the other parts of the apparatus through about 3 feet of 0.003-in. wall, 1/32-in. o.d. inconel and several inches of 0.003-in. wall, 1/64-in. o.d. stainless steel tubing, the latter being nearest the bulbs to give the highest possible thermal resistance between the bulbs and the cryostat. D is a glass U-tube, the lower part of which is a capillary of about 1-mm i.d. A drop of pump oil in the bottom of the U-tube gives a sensitive indication of pressure difference. B is a 3/4-in. o.d. bronze bellows of 25 convolutions, and its volume is adjustable by means of a micrometer (see Fig. 5). A is a fixed volume of such size that with the micrometer set at zero, the total volumes on each side of the oil drop (about 5 cm³) are equal. The system is filled with helium to the desired pressure, (one atmosphere when the whole system is at room temperature) through the valve G, valve H being open to allow the oil drop to come to rest at the bottom of the U-tube.

In operation, one begins with the oil drop at the bottom of the U-tube, the micrometer reading about zero, the thermometer bulbs at the same temperature T, and both valves closed. An increase of the temperature of the left-hand bulb by an amount δT will then cause the oil drop to move to the left. One compensates for this by turning the micrometer, thus changing the volume of the bellows by an amount δV. Application of the perfect gas laws then gives δT:

$$\delta T = \frac{T}{(T_0 V/T\delta V)-1} \cong \frac{T^2\delta V}{T_0 V}, \quad \delta T \ll T, \qquad (7)$$

where T_0 is room temperature. The bulb volume V is constant for all measurements and need be precisely measured only once. Room temperature T_0 can be

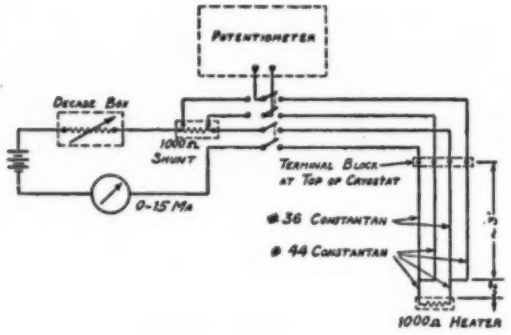


FIG. 6. Heater circuit.

measured with high accuracy with an ordinary mercury thermometer, and the cryostat temperature T can be measured by any of several common methods. As mentioned previously, the cryostat temperature was usually obtained from a measurement of the vapor pressure of the cryostat liquid. The only calibration required is that of the bellows-micrometer combination. In the case of this particular apparatus it was found that

$$\delta V/R=0.200 \text{ cm}^3/\text{mm}, \quad (8)$$

where R is the reading of the micrometer. This calibration was obtained|| by filling the bellows with water and connecting it to a graduated glass tube; the rise of water in the tube as the bellows was compressed gave the change of volume. No observable departure from linearity was noted over the 10-mm range of the micrometer, although the change of volume could be estimated to 0.002 cm³ on the graduated glass tube.

It will be observed from the approximate form of Eq. (7) that δT depends directly on $1/T_0$ and on T^2, so that a small change of T or T_0 during a measurement gives only a small error in the determination of δT, and also so that the "dead" volume of the tubes and bellows does not enter into the determination at all. This latter is perhaps the most attractive feature of the apparatus.

In the derivation of (7) it was assumed that the cold-bulb temperature is constant and equal to the cryostat temperature T. Actually it is more correct to suppose that the cold-bulb temperature during a measurement is slightly greater than T and that δT as computed by Eq. (7) is too small by an amount e given by

$$e=2T\delta Vt/T_0V, \text{ approximately, or } e/\delta T=2t/T, \quad (9)$$

where t is the increase in the cold-bulb temperature during a measurement. The value of K given by Eq. (6) will therefore be too large by the factor $(1+2t/T)$, approximately. We assume that t is proportional to the heater power W, so that one should observe experimentally a small linear dependence of K on W, the true thermal conductivity being obtained by extrapolating the data back to zero heater power.

It is necessary to keep the volume A at precisely the same temperature as the bellows, since a change in temperature of one relative to the other during a measurement will cause error due to motion of the oil drop. For this reason the volume A was kept in good thermal contact with the brass block housing the bellows. It is also essential that the upper parts of the U-tube have considerably greater diameter (3–4 mm i.d.) than the capillary; otherwise, there is danger of a sudden or accidental temperature fluctuation driving oil into the small inconel connecting tubes.

Absolute accuracy of the apparatus as described here is estimated to be about 5 percent or better when $\delta T/T$ is of the order of 0.02. This was sufficiently good for the problem at hand, but there is reason to believe that the accuracy and sensitivity could be increased considerably by a number of refinements, e.g., a smaller capillary in the U-tube, a longer bellows of smaller diameter, correction for deviations from the perfect gas law, and more precise control of T and T_0. At most temperatures, response time of the thermometer for a suddenly applied δT is less than one minute. Response time increases with decreasing temperature, however, and in the liquid helium range is of the order of several minutes, owing to the decreased gas pressure in the thermometer and the rather long and narrow connecting tubes.

To conclude, this apparatus, though not of high precision, has the merit of extreme simplicity in every respect, as compared to other systems capable of similar precision over the whole interval from liquid helium to liquid nitrogen temperatures.

Heater and Heater Circuit

The electrical circuit used to supply power to one end of the specimen under test is given in Fig. 6 and is

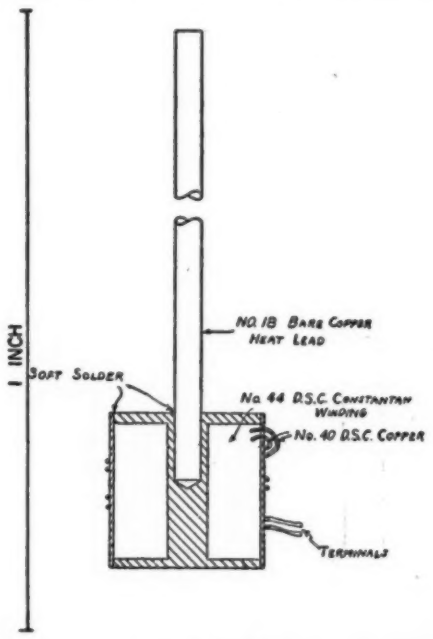


FIG. 7. Heater.

|| By Mr. C. T. Linder.

practically self-explanatory. Constantan current and potential leads were used inside the cryostat to minimize heat conduction down the leads. The emf's that were measured by the potentiometer were of such magnitude that stray thermal emf's were not troublesome.

The heater itself was of special manufacture and is shown in cross section in Fig. 7.

PREPARATION OF SPECIMENS

Most of the commercial alloy specimens used in this work were taken from 1/8-in. o.d. stock tubing or rod and no further preparation was necessary. Monel and inconel tubing** specimens designated in this report as "cold-worked" or "hard-drawn" were from stock bearing the manufacturer's label "Temper No. 3," and those designated as "annealed" from stock labeled "Temper No. 1" or "annealed." The monel rod specimen was from cold-rolled unannealed stock. Nothing is known of the stainless steel specimens except the type numbers 303 and 347, respectively.

Three specimens of the alloy $Cu_{90}Ni_{10}$ were taken from a single forged bar of the alloy, furnished by the Westinghouse Electric Corporation. The treatment of these specimens is detailed in the next section. A fourth specimen was made by remelting a piece of the same forged bar in a graphite crucible in a vacuum furnace, the melt being solidified slowly by lowering the crucible through the furnace at a rate of 3 inches per hour. This furnace and the method of using it have been described by Siegel.[9] Great care was exercised in handling this specimen to keep strain to a minimum. It is estimated that no time was a torque greater than 0.1 inch-pound exerted on any cross section.

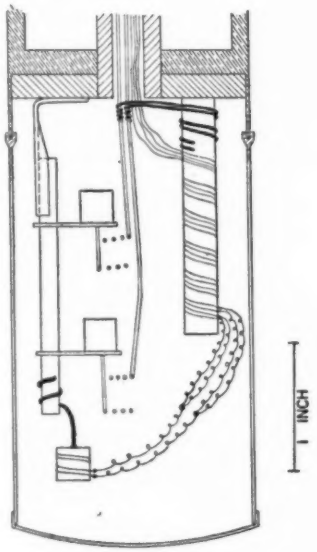


FIG. 8. Mounting of specimen.

** Manufactured by the Superior Tube Company, Norristown, Pennsylvania. Technical information on these alloys is available in Technical Bulletin T-13, *Nickel and Nickel-Base Alloys* (International Nickel Company).

[9] S. Siegel, Phys. Rev. 57, 537 (1949).

TABLE I. Specimen M1. Hard-drawn monel tubing 0.125-in. d, 0.006-in. wall.

T °K	δT °K	K mw/cm-deg	$r\times10^6$ ohm-cm	$Kr/T\times10^8$ w-ohm/deg^2
4.2	0.112	5.26	30	3.7
63.3	12.2	137	...	6.2
77.3	11.45	146	32	5.5
77.3	10.9	153	...	5.9
4.2	0.178	5.33	30	3.7
14.0	0.377	26.4	...	5.6
20.4	0.492	41.2	...	6.0
73.7	3.57	138	32	5.7
77.0	3.71	125	...	4.5
77.0	2.29	141	...	5.8
295	...	...	49	...

Figure 8 is an enlarged view of the working space of the cryostat (labeled *A* in Figs. 2 and 3) showing placement of specimen, thermometer bulbs (*E* and *F* in Fig. 4), and heater (see Figs. 6 and 7). The stainless steel thermometer tubes were thermally grounded to the copper base plate of the helium bomb by means of small copper wires. The four heater leads were wrapped around a copper post which was screwed to the base plate. A brush of fine copper wires (not shown) was soldered to the thermometer tubes at the top end of the vacuum can in order to ground the tubes at this point and prevent heat from being conducted directly from the room temperature portion of the cryostat to the helium bomb. It was estimated *a priori* and confirmed by experiment that the heat conductance of the thermometer tubes and heater leads between the specimen and the helium bomb was no more than 1 percent of the conductance of any specimen.

Thermal contact between the test specimen and the other components of the test set-up was achieved by inserting and soldering the specimen into holes drilled in copper strips 1/32-in. thick, the copper strips being soldered to the cryostat base plate and to the thermometer bulbs, respectively. The lower end of the specimen was soldered to the heat lead from the heater.

RESULTS

Commercial Alloys

Measurements were first carried out on several alloys likely to be of interest in design of low temperature equipment. These were (1) three specimens of monel, one of hard-drawn tubing, one of annealed tubing, and one of hard-drawn solid rod; (2) three specimens of inconel, one of hard-drawn tubing, one of annealed tubing, and one of hot-rolled solid rod; and (3) one specimen of type 303 stainless steel and one of columbium-stabilized type 347 stainless steel furnished by Dr. Horn of General Electric Company. These specimens will be designated by the symbols M1, M2, M3, I1, I2, I3, S1, and S2, respectively. It is probable that all samples came from different melts.††

†† M1 and I1 were samples from the first joint order for special materials distributed by Carnegie Institute of Technology in 1948, and M2 and I2 from the second.

TABLE II. Specimen M2. Annealed monel tubing, 0.126-in.×0.0105-in. wall.

T °K	δT °K	K mw/cm-deg	$r\times10^6$ ohm-cm	$Kr/T\times10^8$ w-ohm/deg²
2.55	0.077	3.69	30.2	4.3
4.25	0.16	9.11	···	6.4
14.0	0.45	46.4	···	9.8
20.5	0.37	74.5	···	10.9
54.0	1.28	140	···	···
63.0	1.80	144	···	7.2
63.7	3.00	147	···	7.2
68.4	2.77	164	···	7.5
77.0	2.71	168	32.1	6.9
295	···	···	50.3	···

Collected results for the monel samples M1, M2, and M3 are given in Tables I, II, and III. M1 was the first of all specimens to be measured, and equipment difficulties as well as lack of practice in using the equipment causes the data on this sample to be somewhat less reliable than later data on other samples. In particular, the values of δT were too large compared to T in the liquid nitrogen region, and the assumed value of T

TABLE III. Specimen M3. Hard-drawn monel rod 0.125-in. d.

T °K	δT °K	K mw/cm-deg	$r\times10^6$ ohm-cm	$Kr/T\times10^8$ w-ohm/deg²
2.6	0.0654	2.16	···	2.35
2.7	0.0510	2.18	···	2.27
4.2	0.14	4.5	28.6	3.00
10.1	0.16	13.1	···	3.7
20.4	0.70	40.6	···	5.6
20.6	0.44	42.8	···	5.9
51.6	1.7	85.5	···	···
59.0	2.8	112	···	···
63.3	2.7	111	···	5.3
75.5	2.3	128	···	5.3
77.0	2.15	148	32.1	6.0
295	···	···	49.2	···

may be in error by a few percent. This objection applies also to measurements on I1, which was chronologically the second specimen. However, it will be noted that the data for M1 and M3 are in fairly good agreement, as might be expected from the similar heat treatment. M2, the annealed specimen, has considerably greater thermal conductivity but practically the same electrical conductivity over the temperature range of measure-

TABLE IV. Specimen I1. Hard-drawn inconel tubing ⅛-in.×0.0105-in. wall.

T °K	δT °K	K mw/cm-deg	$r\times10^6$ ohm-cm	$Kr/T\times10^8$ w-ohm/deg²
14.0	0.45	14.3	···	10.1
20.4	0.64	24.8	90.5	11.0
63.3	4.0	74.4	···	10.3
73.4	4.8	86.6	···	10.3
77.0	3.0	91.0	90.6	10.5
2.6	0.060	1.08	···	3.7
4.2	0.240	3.06	···	6.4
4.2	0.129	2.65	···	5.7
9.6	0.44	8.7	···	8.0
112	···	···	91.0	···
295	···	···	93.7	···

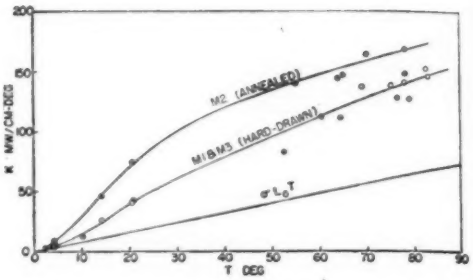


FIG. 9. Thermal conductivity of monel.

ment than M1 or M3; this important fact will be considered in the next section. One also observes that in the case of M1 and M3 L' approaches fairly closely the theoretical low temperature value L_0 in the liquid helium region, or in other words K in this temperature range is approximately equal to $\sigma L_0 T$. The data on monel have been plotted (see Fig. 9) along with the line $\sigma L_0 T$ which is practically the same for all three specimens.

TABLE V. Specimen I2. Annealed inconel tubing 0.126-in.×0.0105-in. wall.

T °K	δT °K	K mw/cm-deg	$r\times10^6$ ohm-cm	$Kr/T\times10^8$ w-ohm/deg²
2.55	0.078	2.00	···	7.9
4.25	0.19	4.83	102.2	11.3
14.0	0.51	27.0	···	19
20.5	0.48	40.7	···	20
20.5	0.73	41.5	···	20
22.1	0.80	47.0	···	21
54.0	2.3	96.3	···	17.5
63.3	3.2	100	···	15
63.7	5.0	104	···	16
68.4	3.1	96.6	···	14
77.0	4.7	110	100.3	14
295	···	···	102.8	···

Most of the preceding remarks about monel apply to inconel as well. Specimen I1 in this case is hard-drawn, I2 and I3 being more or less annealed. The effect of

TABLE VI. Specimen I3. Hot-rolled inconel rod ½-in. d.

T °K	δT °K	K mw/cm-deg	r*
2.59	0.06	1.96	
4.25	0.16	4.89	
14.0	0.48	27.7	
58.0	2.4	91	
63.0	4.1	93.5	
63.3	2.3	93.5	
63.3	4.0	96.2	
70.6	4.0	96.5	
76.2	3.8	100	
2.59	0.045	2.13	
4.25	0.20	5.17	
4.25	0.13	4.93	
10.1	0.20	16.8	
14.0	0.46	29.7	
58.7	2.2	99.7	
59.2	2.2	99.7	
77.8	2.0	107	

* Not measured.

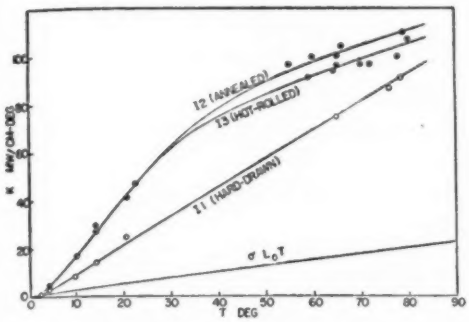


FIG. 10. Thermal conductivity of inconel.

annealing on thermal conductivity and on L' is again evident from the data in Tables IV, V, and VI, and the graph in Fig. 10. L' decreases with temperature below 20°K but even for I1 remains somewhat above the theoretical value at the low temperature limit.

TABLE VII. Specimen S1. Type 303 stainless steel ⅛-in. d.

T °K	δT °K	K mw/cm-deg	$r\times10^{6}$ ohm-cm	$Kr/T\times10^{8}$ w-ohm/deg²
4.2	0.108	2.47	51.2	3.0
4.2	0.098	2.53	···	3.0
10.1	0.38	6.63	···	3.3
20.4	0.079	19.9	···	4.9
20.6	1.04	22.0	···	5.4
56.9	2.9	66.9	···	5.9
58.0	2.8	65.7	···	5.8
59.7	2.8	66.7	···	5.6
63.3	2.6	71.9	···	5.7
66.8	5.5	73.3	···	5.5
77.0	2.15	84.3	51.9	5.6
2.59	0.059	1.25	···	2.5
4.25	0.15	2.40	···	2.85
10.1	0.46	6.93	···	3.4
19.4	1.2	19.8	···	5.2
58.3	3.2	66.4	···	5.7
59.5	3.1	67.2	···	5.7
77.8	2.6	82.3	···	5.4
295	···	···	73.6	···

Of all the commercial alloys, specimens S1 and S2 showed the most nearly linear dependence of K on T, and also the smallest departure of L'/L_0 from unity, as can be seen from the data in Tables VII and VIII and the graphs in Figs. 11 and 12.

TABLE VIII. Specimen S2. Type 347 Cb-stabilized stainless steel ⅛-in. d.

T °K	δT °K	K mw/cm-deg	$r\times10^{6}$ ohm-cm	$Kr/T\times10^{8}$ w-ohm/deg²
4.25	0.17	2.34	52.7	2.8
4.25	0.18	2.34	···	2.8
14.0	0.52	11.7	···	4.3
58.5	3.5	63.8	···	5.6
63.2	3.4	66.0	···	5.4
63.3	3.4	67.0	···	5.5
70.7	3.1	73.2	···	5.4
76.2	2.9	77.2	53.4	5.3
295	···	···	74.4	···

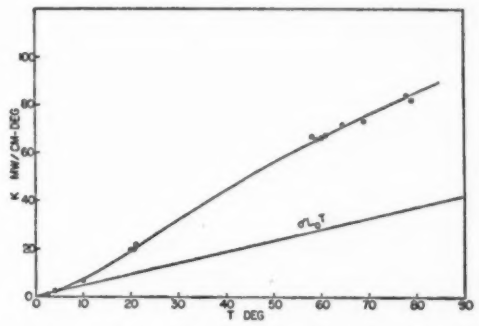


FIG. 11. Thermal conductivity of stainless steel (type 303).

Copper-Nickel Alloys

The consistent differences in L'/L_0 between cold-worked and annealed monel and inconel suggested a series of measurements on specimens varying only in degree of cold-work and grain size, all taken from a single sample of alloy. Without too much regard for precise composition, the alloy chosen was 90 percent copper, 10 percent nickel. Copper-nickel alloys of about this composition have two inherent properties worth noting: (1) Copper and nickel have very nearly identical lattice properties, so that as far as purely lattice properties are concerned, the alloy should behave the same as pure copper, and (2) the d-band of energy levels is thought to be completely filled so that the band structure is the same as that of copper, a fact of some importance in theoretical consideration of electronic properties. In fact, this alloy should behave very much like copper except for the randomly distributed irregularities in the otherwise periodic potential of the free electrons. These irregularities strictly limit the electronic mean free path at low temperatures, giving according to measurement a residual resistivity $1/\sigma_0$ only about 15 percent smaller than the room temperature resistivity.

Data were taken on four specimens of this alloy, which will be designated as CN1, CN2, CN3, and CN4. CN1 was taken from the original forged bar and machined to size with no other treatment. CN2 was cold-worked by rolling down a thickness of 0.25-in. to 0.14-in. before being machined to size. CN3 was rolled from a 1/2-in. square cross section to 0.22-in. by 0.24-in. before machining. The effect of this cold-working on the low temperature thermal conductivity is clearly evident

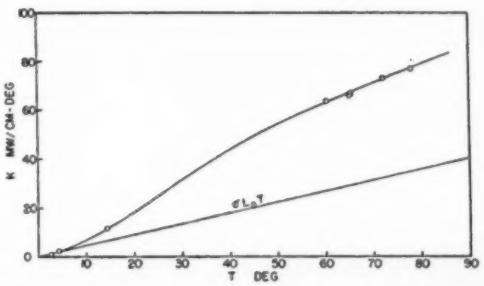


FIG. 12. Thermal conductivity of stainless steel (type 347).

TABLE IX. Specimen CN1. 90% Cu 10% Ni (annealed) ⅛-in. d.

T °K	K mw/cm-deg	$r\times10^3$ ohm-cm	$Kr/T\times10^3$ w-ohm/deg²
3.33	7.96	...	2.99
4.21	12.0	...	3.56
14.1	96.4	...	8.5
16.7	123	...	9.2
19.7	157	12.50	10.0
64.0	388	...	7.7
78.0	378	12.72	6.2
296	...	14.68	...

from Tables IX, X, and XI and the graph in Fig. 13. CN4 was prepared by the single crystal growing technique described in the preceding section. Data for this sample are given in Table XII and are plotted also in Fig. 13. Electrical resistivities of the first three specimens are very nearly equal, whereas that of CN4 is slightly higher. This was rather unexpected, and so far a satisfactory explanation has not been found. It is probably that CN4 was somewhat inhomogeneous in

TABLE X. Specimen CN2. 90% Cu 10% Ni (cold-worked) ⅛-in. d.

T °K	K mw/cm-deg	$r\times10^3$ ohm-cm	$Kr/T\times10^3$ w-ohm/deg²
3.03	5.80	...	2.40
3.61	7.34	...	2.54
4.21	9.17	...	2.72
4.40	10.3	...	2.93
14.07	60.1	...	5.3
17.0	82.5	...	6.1
19.7	104.0	...	6.6
71.3	343	...	6.1
76.2	358	12.65	5.95
296	...	14.69	...

composition, being richer in nickel toward the end where solidification occurred first. If this is true, then the conductivity data represents a sort of average over composition and is not necessarily correct for the average composition; the error, however, would probably not be greater than one or two percent. In any event, it can be shown that macroscopic inhomogeneity along the length of the specimen would decrease rather than increase the measured resistivity.

TABLE XI. Specimen CN3.
90% Cu 10% Ni (severely cold-worked) ⅛-in. d.

T °K	K mw/cm-deg	$r\times10^3$ ohm-cm	$Kr/T\times10^3$ w-ohm/deg²
3.63	7.35	...	2.53
4.20	8.92	...	2.65
4.72	10.8	...	2.86
14.1	58.7	...	5.2
17.2	78.6	...	5.7
19.4	98.6	...	6.4
64.0	326	...	6.4
70.6	333	...	5.9
78.7	348	12.63	5.6
298	...	14.65	...

DISCUSSION OF RESULTS

Practical Considerations

Our thermal conductivity data for two types of stainless steel agree within a few percent with hydrogen temperature data obtained by Wilkinson and Wilks[10] for a sample of unspecified type, and also with nitrogen temperature data taken by Mr. David Lyons.‡‡ The random nature of this selection of samples encourages the view that these results are of fairly general validity. Thermal and electrical conductivities of monel and inconel at low temperatures apparently have not been published. Out data for hard-drawn inconel are comparable with those obtained by Karweil and Schafer[11] for "contracid," an alloy of similar composition but with the addition of several percent molybdenum. From a practical point of view it is to be noted that both stainless steel and hard-drawn inconel are about as good thermal insulators in the low temperature region as any alloys yet investigated. However, the only extensive series of measurements yet published are those of

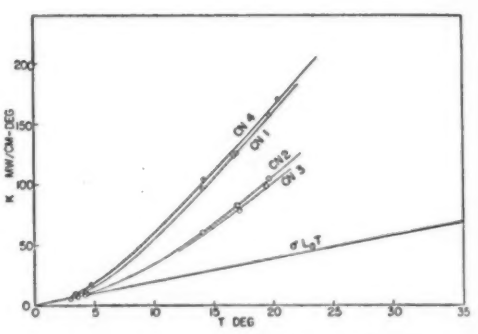


FIG. 13. Thermal conductivity of $Cu_{90}Ni_{10}$.

Karweil and Schafter on contracid, German silver, iron, and steel, so the field is open for further investigation.

The practical importance of cold work in reducing heat conduction in alloys has been the subject of speculation but not of quantitative study up to this time.§§ Our results on monel, inconel, and $Cu_{90}Ni_{10}$ show that at low temperatures heat resistance of alloys can be increased by a factor of 2 or more by cold-working, a fact of considerable significance in the design of cryogenic equipment.

Comparison with Theory

A general survey of experimental results show that L'/L_0 has a broad maximum considerably larger than unity in the liquid nitrogen region or above. L'/L_0 greater than one is a theoretical indication of appreciable lattice heat conduction; the most convincing argument for the presence of lattice conduction is, however, the

[10] K. R. Wilkinson and J. Wilks, J. Sci. Instr. **26**, 19 (1949).
‡‡ Private communication.
[11] J. Karweil and K. Schafer. Ann. Physik. **36**, 567 (1939).
§§ See, however, A. Eucken and K. Dittrich, Z. Phys. Chem. **125**, 211 (1927).

TABLE XII. Specimen CN4.
90% Cu 10% Ni (single crystal) ⅛-in. d.

T °K	K mw/cm-deg	$r\times10^6$ ohm-cm	$Kr/T\times10^8$ w-ohm/deg²
3.45	9.80	···	3.70
4.80	17.3	···	4.69
14.2	103.6	···	9.5
17.0	135	···	10.3
20.5	170	13.0	10.8
64.5	350	···	7.1
73.0	361	···	6.4
79.3	357	13.10	5.7
298	···	15.04	···

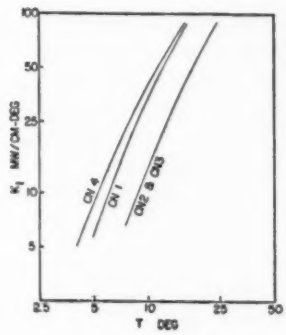


FIG. 14. Lattice thermal conductivity of $Cu_{90}Ni_{10}$.

observed disproportionate effect of cold-work on thermal and electrical conductivities; namely, that while thermal conductivity may be decreased by a factor of 1.5 or 2 by cold-work, electrical conductivity is not measurably affected. According to Eq. (2), any effect of lattice dislocations on the electronic mean free path should change the thermal and the electrical conductivities proportionately and leave the ratio unchanged. Since the theory has been developed only for the case of free electrons and partially filled D-bands of energy levels, it might be proposed that the large values of L' could be explained on the basis that the electrons are not perfectly free, as has been suggested by Seitz[13] for the case of beryllium and β-manganese. However, this is inconsistent with the fact that no order of magnitude difference is noted between the L' values for the commercial alloys and for the $Cu_{90}Ni_{10}$, in which alloy the D-band is completely or nearly completely filled.|| ||

It might be expected on the basis of theory that sufficient cold-working would reduce the lattice conductivity to zero and leave only the electronic component which they obey the Wiedemann-Franz law within a few percent, or at any rate give $L'/L_0<1$. This expectation is not borne out by the results on CN2 and CN3, which indicate that L'/L_0 approaches a constant value considerably greater than one with increasing cold-work. This means that cold-working alone does not produce enough lattice defects to effectively eliminate lattice conduction, except perhaps at liquid helium temperatures where L' for the cold-worked alloys and the stainless steels does approach L_0 quite closely.

Theory predicts that in the neighborhood of absolute zero the lattice conductivity K_l should disappear leaving $K=K_e=\sigma L_0 T$, and generally speaking, this is what is observed, although in the annealed specimens lattice conduction is still comparable to electronic conduction at the lowest temperature obtainable with the equipment used. In no case do we observe behavior of the type reported by Allen and Mendoza,[14] who found that the thermal conductivity of a German silver specimen was proportional to T^3 and was much less than $\sigma_0 L_0 T$ in the liquid helium range.

The behavior of the lattice conductivity of the copper-nickel specimens, computed by means of Eqs. (1) and (2) and the curves of Fig. 13, is shown in Fig. 14. Between 5 and 15°K, K_l for CN4 may be represented within a few percent by the equation

$$K_l=0.39T^2 \text{ mw/cm-deg.} \tag{10}$$

For comparison, we have from a paper published recently by Hulm[15]

$$K_l=0.22T^{2.0} \text{ ms/cm-deg} \tag{11}$$

for an annealed (?) sample of the alloy $Cu_{80}Ni_{20}$. Finally, for the alloy $Cu_{70}Ni_{30}$ we can compute K_l from the thermal conductivity data of Wilkinson and Wilks[10] and electrical conductivity given by Landolt and Bornstein.[16] The result is, very roughly,

$$K_l=0.11T^2 \text{ mw/cm-deg,} \tag{12}$$

correct in the neighborhood of 12°K. Comparison of these three equations shows that the lattice conductivity of copper-nickel alloys is roughly inversely proportional to the nickel content in this concentration range. As a consequence of this we should expect the lattice conductibility of pure copper to be larger than that of, say, $Cu_{90}Ni_{10}$ by at least an order of magnitude, and perhaps comparable to that of insulating crystalline solids.

Between 10 and 20°K, K_l for CN2 and CN3 is slightly more than 1/3 the lattice conductivity of CN1. We feel justified in assuming, therefore, that in these two specimens the lattice waves are scattered almost entirely by dislocations resulting from cold-working. If we further make the approximation that the mean free path for this type of scattering is independent of frequency, then Eq. (3) may be written

$$K_l=3Nvk\frac{T^3}{\theta}(25.9\ldots). \tag{13}$$

[13] F. Seitz, *Modern Theory of Solids* (McGraw-Hill Book Company, Inc., New York, 1940), p. 178.

|| This does little to alter the possible validity of Seitz's suggestion. The anomalous behavior of beryllium takes place at high temperatures and is of quite a different nature. The experimental results for β-manganese consist of two points in the neighborhood of 80°K.

[14] J. F. Allen and E. Mendoza, Proc. Cambridge Phil. Soc. **44**, 280 (1948).

[15] J. K. Hulm, Proc. Phys. Soc. (London) **64**, 207 (1951).

[16] Landolt and Bornstein, *Physikalisch-Chemische Tabellen* (Verlag Julius Springer, Berlin, 1923), Hw, p. 1054.

Use of this equation for an order of magnitude calculations of l gives 2×10^{-5} cm. The same result can be obtained from Eq. (4). This is considerably greater than the distance between slip bands in cold-worked copper,[17] and indicates that the reflectivity of the slip bands for acoustic waves is of the order of 0.1 or less. Since the wavelength of maximum amplitude in the vibration spectrum of the Debye continuum corresponding to copper at a temperature of $\theta/10$ is approximately $2.5h\nu/k\theta$ or about 10 angstrom units, the logarithmic decrement per cycle for thermal vibrations at this temperature turns out to be of the order of 0.005.

If we assume that the logarithmic decrement rather than the mean free path is independent of frequency, then it follows that the mean free path varies inversely as the frequency, and in the region where C_v varies as T^3 the lattice conductivity must vary as T^2. Such temperature dependence for K_l actually corresponds somewhat more closely to experimental results than Eq. (13), derived on the basis on constant mean free path.

[17] F. Seitz, "Theory of plastic flow in single crystals," Symposium on the Plastic Deformation of Crystalline Solids, Mellon Institute, Pittsburgh, Pennsylvania (May, 1950).

ACKNOWLEDGMENT

The writers are indebted to Professor J. E. Goldman and Professor R. Becker for helpful discussions on various theoretical topics, and to Frederick Darnell and Edward Lloyd for assistance in making measurements. They also wish to thank the Research Corporation for a grant-in-aid, from which some of the equipment used in this work was purchased.

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 5 MAY, 1952

The Radiation Pattern of an Antenna over a Circular Ground Screen*

JAMES E. STORER
Cruft Laboratory, Harvard University, Cambridge, Massachusetts
(Received February 7, 1952)

In a previous paper [*The Input Impedance of an Antenna Over a Finite Ground Screen*, Technical Report No. 119 (Cruft Laboratory, Harvard University, November 30, 1950)] the author treated the problem of the input impedance of an antenna erected over a large circular screen. Using the formulation developed therein, an expression is obtained for the entire radiation pattern of an antenna over a large circular screen.

The most important difference between the pattern of an antenna over a large screen and an antenna over an infinite screen is the presence of lobes over the antenna. From the expression for the radiation pattern, simple formulas are obtained for: (a) the number of lobes, (b) the angular position of the lobes, and (c) the angle, β_{max}, within which the lobes occur. One important result obtained is that

$$\theta_{max}\sim(1/d^{\frac{1}{2}}),$$

where d is the diameter of the screen. It is apparent that the effects of a finite screen decrease very slowly as the diameter of the screen is increased.

I. INTRODUCTION

IN many antenna experiments, the antenna is placed over a large conducting screen, as indicated in Fig. 1. The assumption is then made that the experimental results are in substantial agreement with those that would be obtained if the screen were of infinite size. Recent experimental work by Meier and Summers,[1] however, indicates that the finiteness of the screen may have appreciable effect on the measurements. Hence, it is worth while to examine this problem theoretically and derive expressions for the effects produced by the use of finite screens.

In a previous paper,[2] the author obtained an expression for the input impedance of an antenna erected over a large circular screen. In the present paper the

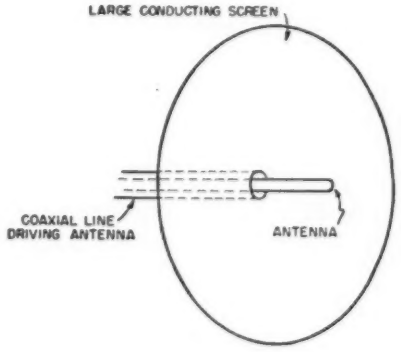


FIG. 1. Typical experimental set-up for performing measurements on antennas.

* The research reported in this paper was carried out under a Fellowship awarded the author by the AEC. Publication as Technical Report No. 126 was made possible through support extended Cruft Laboratory, Harvard University, jointly by the Navy Department (ONR), the Signal Corps of the U. S. Army, and the U. S. Air Force, under ONR Contract N5-ori-76, T. O. 1.

[1] A. S. Meier and W. P. Summers, *Measured Impedances of Vertical Antennas over Finite Ground Planes*, Report 233-3, (Antenna Laboratory, Ohio State Research Foundation, October, 1946).

[2] J. E. Storer, J. Appl. Phys. 22, 1058 (1951).

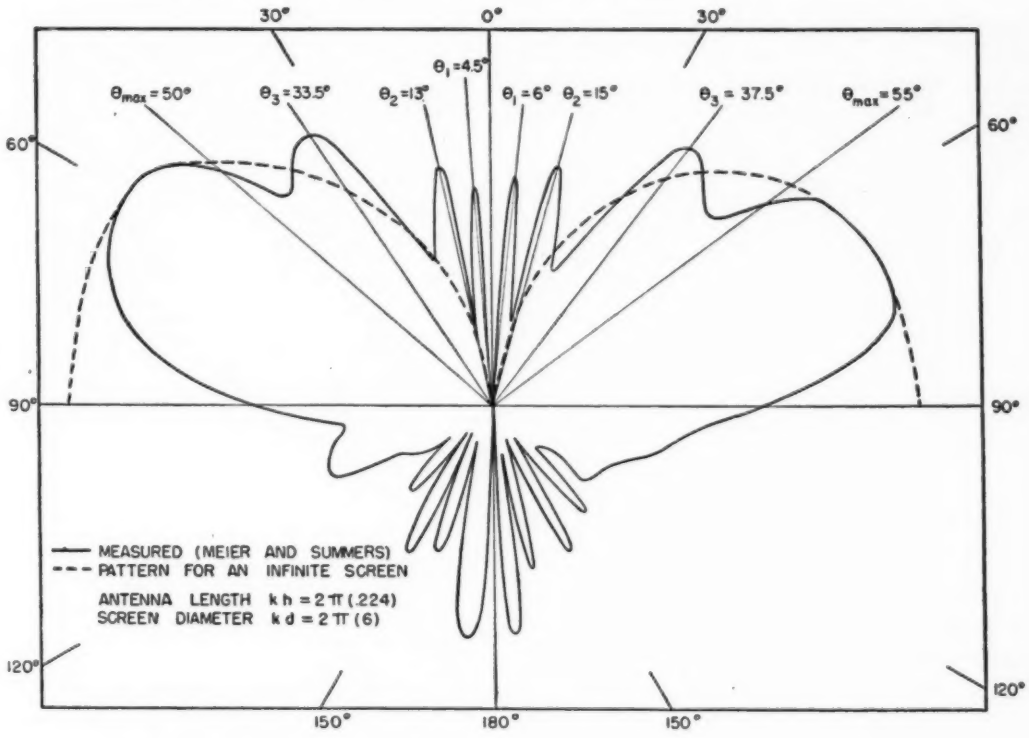


FIG. 2. The radiation pattern of an antenna over a finite ground screen.

complementary problem of the radiation pattern will be discussed. To the best of the author's knowledge, the only research on this subject to date is contained in a paper by Leitner and Spence.[3] Their solution for the radiation pattern is expressed in terms of a series of spheroidal functions. However, for a large screen their solution is of limited usefulness because of the lengthy numerical computations necessary for evaluation of the series. Moreover, their work is confined to the particular case of a quarter-wavelength antenna. By using a variational formulation, this paper will obtain, for large screens, simple expressions for the change of the radiation pattern as a function of the diameter of the screen. There will be no restriction placed on the length of the antenna.

Before entering into a mathematical discussion, it is useful to examine a pattern measured experimentally. Figure 2 contains a typical pattern measured by Meier and Summers[1] with the pattern for an infinite screen superimposed on it. While the presence of fields to the rear of the screen is to be expected, the occurrence of lobes over the antenna is a considerably more interesting change.

From an experimental point of view, the pattern to the side and rear of the screen is of little interest since it generally cannot be measured. Moreover, if it were measured, the presence of the driving apparatus and the screen supports would make the prediction of this part of the pattern impossible. In the region $25°<\theta<75°$, the pattern of a large screen is so nearly identical to the pattern of an infinite screen that the differences are undetectable experimentally. (This is not true of the pattern in Fig. 2 since the screen is not large in the sense that $kd\gg1$.) Therefore, while it is possible to obtain theoretically the entire pattern over a large screen, experimental comparisons are in reality limited to the region of small angles where the lobes occur. Hence, while a general formula will be developed for the radiation pattern, it will be simplified and discussed only for small angles.

II. THE RADIATION PATTERN

Using a cylindrical coordinate system (r, z, ϕ), oriented so that the circular ground screen coincides with

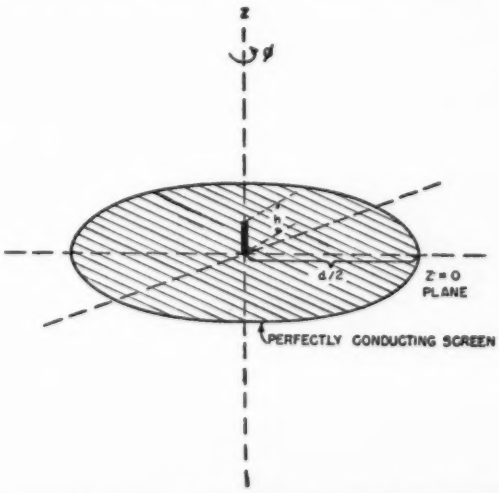


FIG. 3. Orientation of the cylindrical coordinate system.

[3] A. Leitner and R. D. Spence, J. Appl. Phys. 21, 1001 (1950).

the plane $r=0$ and the antenna lies along the z axis (see Fig. 3), it is apparent that circular symmetry exists. Therefore the only electromagnetic fields present are H_ϕ, E_r, E_z. It can be shown [reference 2, Eq. (7)] that H_ϕ may be written in the form,

$$H_\phi=-\frac{1}{4\pi}e^{-i\omega t}\frac{\partial}{\partial r}\left\{\int_0^h\left[\frac{\exp\{ik[(z-z')^2+r^2]^{\frac{1}{2}}\}}{[(z-z')^2+r^2]^{\frac{1}{2}}}+\frac{\exp\{ik[(z+z')^2+r^2]^{\frac{1}{2}}\}}{[(z+z')^2+r^2]^{\frac{1}{2}}}\right]I(z')dz'\right\}$$

$$-\frac{2ik}{\zeta_0}e^{-i\omega t}\int_{d/2}^{\infty}G(z,r,r')\xi(r')r'dr',\quad z>0 \qquad (1a)$$

$$H_\phi=\frac{2ik}{\zeta_0}e^{-i\omega t}\int_{d/2}^{\infty}G(z,r,r')\xi(r')r'dr',\quad z<0 \qquad (1b)$$

where d denotes the diameter of the screen, h denotes the height of the antenna, $k=2\pi/\lambda$, where λ is the free-space wavelength, $\zeta_0=120\pi$ ohms, $I(z)$ is the total current on the antenna, $\xi(r)=[E_r]_{z=0}$, and

$$G(z,r,r')=\frac{1}{4\pi}\int_0^{2\pi}\frac{\exp\{ik[z^2+r^2+r'^2-2rr'\cos\phi]^{\frac{1}{2}}\}}{[z^2+r^2+r'^2-2rr'\cos\phi]^{\frac{1}{2}}}\cos\phi d\phi.$$

The part of H_ϕ involving an integral of $I(z')$ represents the fields that would be present if the screen were of infinite size; the remaining integrals involving $\xi(r')$ represent the changes introduced by the use of a finite screen.

In discussing the far fields, a spherical coordinate system (R,θ,ϕ), related to the cylindrical coordinate system by

$$z=R\cos\theta,\quad r=R\sin\theta,\quad \phi=\phi,$$

will be used. Since there is no dependence of the fields on the ϕ-coordinate, it is necessary to plot the fields only as a function of θ. The far zone is defined by the inequalities

$$kR\gg 1,\quad R^2\gg d^2. \qquad (2)$$

Using these inequalities, and denoting far fields by the subscript ff, $H_{\phi ff}$ may be determined by taking the limiting form of Eq. (1). Thus, using (2), it is seen that

$$\frac{\exp\{ik[z^2+r^2+r'^2-2rr'\cos\phi]^{\frac{1}{2}}\}}{[z^2+r^2+r'^2-2rr'\cos\phi]^{\frac{1}{2}}}\cong\frac{e^{ikR}}{R}e^{-ikr'\sin\theta\cos\phi},$$

and hence

$$G(z,r,r')\cong\frac{e^{ikR}}{4\pi R}\int_0^{2\pi}e^{-ikr'\sin\theta\cos\phi}\cos\phi d\phi.$$

Also,

$$\left[\frac{\exp\{ik[(z-z')^2+r^2]^{\frac{1}{2}}\}}{[(z-z')^2+r^2]^{\frac{1}{2}}}+\frac{\exp\{ik[(z+z')^2+r^2]^{\frac{1}{2}}\}}{[(z+z')^2+r^2]^{\frac{1}{2}}}\right]\cong\frac{e^{ikR}}{R}e^{-ikz'\cos\theta}+\frac{e^{ikR}}{R}e^{ikz'\cos\theta}=2\frac{e^{ikR}}{R}\cos[kz'\cos\theta].$$

Combining these expressions with Eq. (1) yields

$$H_{\phi ff}=\frac{e^{ikR-i\omega t}}{R}\left\{\frac{-ik\sin\theta}{2\pi}\int_0^h\cos[kz'\cos\theta]I(z')dz'\right.$$

$$\left.-\frac{ik}{2\pi\zeta_0}\int_0^{2\pi}\int_{d/2}^{\infty}e^{-ikr'\sin\theta\cos\phi}\xi(r')\cos\phi r'dr'd\phi\right\}z>0 \qquad (3)$$

$$H_{\phi ff}=\frac{e^{ikR-i\omega t}}{R}\left\{\frac{ik}{2\pi\zeta_0}\int_0^{2\pi}\int_{d/2}^{\infty}\times e^{-ikr'\sin\theta\cos\phi}\xi(r')\cos\phi r'dr'd\phi\right\}z<0.$$

Since $\xi(r)$ is a radial component of the electric field, $\xi(r)\sim e^{ikr}/r^2$ for large r. Hence the expression for $H_{\phi ff}$, Eq. (3), is not valid at $\theta=\pi/2$ since the integral involving $\xi(r)$ does not converge at $\theta=\pi/2$. This is not surprising considering the manner in which the original expression for H_ϕ, Eq. (1), was derived. Equation (3), however, is valid for all $\theta\neq\pi/2$. Quite obviously if an incorrect $\xi(r)$ were used to evaluate (3), poor results would be obtained in the region of $\theta\cong\pi/2$.

In the author's first paper[2] it was shown that to a good approximation the function $\xi(r)$ may be written in the form

$$\xi(r)=\frac{e^{ikr}}{[r-d/2]^{\frac{1}{2}}}F(r), \qquad (4)$$

where $F(r)$ is a slowly varying function of r. This approximation may be used in (3) to evaluate $H_{\phi ff}$. As pointed out previously, such a procedure cannot possibly yield good results in the region $\theta\cong\pi/2$, say $85°<\theta<95°$. In order to obtain an expression valid near $\theta=\pi/2$, it is necessary to use (3) together with the integral equation [reference 2, Eq. (8)] that the correct $\xi(r)$ satisfies. However, no attempt will be made to do this since the region $\theta\cong\pi/2$ is not available for experimental comparison.

It is convenient at this time to introduce the restriction that the ground screen be large, i.e.,

$$kd\gg 1,\quad d^2\gg h^2. \qquad (5)$$

The inequality $kd\gg 1$ is interpreted to mean that the screen is at least 10 wavelengths in diameter. With the aid of these inequalities, the approximation for $\xi(r)$, (4), may be inserted into Eq. (3) and the integral evaluated. For details of this process the reader is re-

ferred to the author's original paper.[4] As a result one obtains

$$H_{\phi ff}=\begin{cases}\dfrac{e^{ikR-i\omega t}}{R}\left\{\dfrac{-ik\sin\theta}{2\pi}\displaystyle\int_0^h\cos[kz'\cos\theta]I(z')dz'+\dfrac{ik}{4\pi^2\sqrt{2}}\left[\displaystyle\int_0^h I(z)dz\right]\left[\displaystyle\int_0^{2\pi}\frac{e^{ikd/2\,(1-\sin\theta\cos\phi)}}{[1-\sin\theta\cos\phi]^{\frac{1}{2}}}\cos\phi d\phi\right]\right\} & 0\leq\theta<\dfrac{\pi}{2}\\ \dfrac{e^{ikR-i\omega t}}{R}\left\{-\dfrac{ik}{4\pi^2\sqrt{2}}\left[\displaystyle\int_0^h I(z)dz\right]\left[\displaystyle\int_0^{2\pi}\frac{e^{ikd/2\,(1-\sin\theta\cos\phi)}}{[1-\sin\theta\cos\phi]^{\frac{1}{2}}}\cos\phi d\phi\right]\right\} & \dfrac{\pi}{2}<\theta\leq\pi.\end{cases}\quad(6)$$

It is conventional to express far fields in terms of the quantity actually measured, the *rms* value of $E_{\theta ff}$. Since $E_{\theta ff}{}^{rms}=\zeta_0 H_{\phi ff}{}^{rms}$, (6) can be used to obtain the final form for the radiation pattern of an antenna erected over a large circular screen. It is

$$E_{\theta ff}{}^{rms}=\zeta_0 H_{\phi ff}{}^{rms}=\frac{\zeta_0}{\sqrt{2}}|H_{\phi ff}|=\begin{cases}\left|-ik\theta\displaystyle\int_0^h\cos[kz\cos\theta]I(z)dz\right. & \\ \left.+\dfrac{ik}{2\pi\sqrt{2}}\left[\displaystyle\int_0^h I(z)dz\right]\left[\displaystyle\int_0^{2\pi}\frac{e^{ikd/2\,(1-\sin\theta\cos\phi)}}{[1-\sin\theta\cos\phi]^{\frac{1}{2}}}\cos\phi d\phi\right]\right| & 0\leq\theta<\dfrac{\pi}{2}\\ \left|-\dfrac{ik}{2\pi\sqrt{2}}\left[\displaystyle\int_0^h I(z)dz\right]\left[\displaystyle\int_0^{2\pi}\frac{e^{ikd/2\,(1-\sin\theta\cos\phi)}}{[1-\sin\theta\cos\phi]^{\frac{1}{2}}}\cos\phi d\phi\right]\right| & \dfrac{\pi}{2}<\theta\leq\pi.\end{cases}\quad(7)$$

For values of θ *not* near 0 and π, the ϕ integral could readily be evaluated by a saddle-point integration. However, there is no object in doing this since experimental comparisons are available only for $\theta\cong 0$.

III. THE RADIATION PATTERN FOR SMALL ANGLES

Examination of Eq. (7) indicates that the far fields are nearly identical to those of an infinite screen unless the exponential part of the ϕ integrand is *not* a rapidly oscillating function, i.e., $(kd/2)\sin\theta$ is relatively small. Since the inequality $kd\gg 1$ is intended to mean $kd\geq 60$ (or, $d\geq 10\lambda$) the exponential is not rapidly oscillating only when θ is near 0 and π. This is in agreement with the experimental result that the only significant changes in the radiation pattern are the lobes occurring at small angles (and, of course, the radiation to the rear of the screen). For small θ, e.g., $\theta\leq\frac{1}{2}$ (or $\theta<30°$), Eq. (7) may be simplified by using the approximations

$$\sin\theta\cong\theta,\quad\cos\theta\cong 1,\quad[1-\sin\theta\cos\phi]^{\frac{1}{2}}\cong 1.$$

Hence, for small θ,

$$E_{\theta ff}{}^{rms}\sim\left|-ik\theta\int_0^h\cos kzI(z)dz+\frac{ik}{2\pi\sqrt{2}}\left[\int_0^h I(z)dz\right]\left[\int_0^{2\pi}e^{ikd/2(1-\theta\cos\phi)}\cos\phi d\phi\right]\right|.$$

Using the integral répresentation,

$$J_1(x)=\frac{i}{2\pi}\int_0^{2\pi}e^{-ix\cos\phi}\cos\phi d\phi,$$

this may be simplified to

$$E_{\theta ff}{}^{rms}\sim\left|\underbrace{-ik\theta\int_0^h\cos kzI(z)dz}_{\text{I}}+\underbrace{\frac{k}{\sqrt{2}}\left[\int_0^h I(z)dz\right]e^{i(kd/2)\theta}J_1\left(\frac{kd\theta}{2}\right)}_{\text{II}}\right|.\quad(8)$$

In Eq. (8), part I represents the fields present over an infinite screen, and part II the changes due to the use of a finite screen. It is apparent that the changes of part II become negligible as soon as θ takes on sizable values. It is possible to use (8), as it stands, to predict the structure of the lobes at small angles. However, this detailed structure is not particularly interesting and it is more convenient to use (8) to estimate the following quantities: (a) angular position of the lobes, (b) angular separation of the lobes, (c) number of prominent lobes, and (d) range, $0<\theta<\theta_{\max}$, over which the lobe structure occurs.

It is apparent from Eq. (8) that the minimum points between the lobes occur when, approximately,

$$J_1(kd\theta/2)=0.$$

Also, these minimum points should correspond very closely to the plot for the infinite screen since the effects of a finite screen, II, disappear at these points. The maximum points of the lobes occur when, approximately,

$$J_1(kd\theta/2)\text{ is a maximum.}\quad(9)$$

[4] J. E. Storer, *The Radiation Pattern of an Antenna over a Finite Ground Screen*, Technical Report No. 126 (Cruft Laboratory, Harvard University, June, 1951).

Using the asymptotic formula,

$$J_1(x)\frac{2}{\pi x}\sin\left(x-\frac{\pi}{4}\right),$$

the condition (9) is equivalent to

$$\sin\left(\frac{kd\theta}{2}-\frac{\pi}{4}\right)=\pm 1, \quad \text{or} \quad \frac{kd\theta}{2}-\frac{\pi}{4}=n\pi+\frac{\pi}{2}. \tag{10}$$

The asymptotic formula may be used to good approximation for lobes other than the first. Hence, denoting points of maximum by θ_n, Eqs. (9) and (10) yield

$$\theta_n=\left\{\begin{matrix}\frac{2}{kd}(1.84)\\ \frac{2}{kd}\left(n\pi+\frac{3\pi}{4}\right)\end{matrix}\right\}=\left\{\begin{matrix}0.59\frac{\lambda}{d}\text{ radians} & n=1\\ (n+\frac{3}{4})\frac{\lambda}{d}\text{ radians} & n=2,3,4\cdots.\end{matrix}\right.$$

The angular separation $\Delta\theta$ of the lobes is then

$$\Delta\theta=\theta_{n+1}-\theta_n\cong\frac{\lambda}{d}\text{ radians.}$$

From (8) it is seen that the prominent lobes will have disappeared when

$$\text{Magnitude}\left[-ik\theta\int_0^h \cos kzI(z)dz\right]$$

$$\geq 4\times\text{Magnitude}\left[+\frac{k}{\sqrt{2}}\left\{\int_0^h I(z)dz\right\}e^{i(kd\theta/2)}J_1\left(\frac{kd\theta}{2}\right)\right].$$

Replacing the Bessel function by its asymptotic value,

$$k\theta\left|\int_0^h \cos kzI(z)dz\right|\geq 4\left[k\left(\frac{2}{\pi kd\theta}\right)^{\frac{1}{2}}\left|\int_0^h I(z)dz\right|\right]. \tag{11}$$

Denote by $\theta_{\max}$ the angle within which all prominent lobes occur; it is then seen by solving (11) that

$$\theta_{\max}=\left[\frac{32}{\pi kd}\right]^{\frac{1}{2}}\left|\frac{\int_0^h I(z)dz}{\int_0^h \cos kzI(z)dz}\right|^{\frac{1}{2}}$$

$$=1.175\left(\frac{\lambda}{d}\right)^{\frac{1}{2}}\left|\frac{\int_0^h I(z)dz}{\int_0^h \cos kzI(z)dz}\right|^{\frac{1}{2}}.$$

Denote the number of prominent lobes by N; it is then seen that, approximately,

$$N+1=\frac{\theta_{\max}}{\Delta\theta}=1.175\left(\frac{d}{\lambda}\right)^{\frac{1}{2}}\left|\frac{\int_0^h I(z)dz}{\int_0^h \cos kzI(z)dz}\right|^{\frac{1}{2}}.$$

For antenna lengths such that $kh\leqslant 3\pi/4$, it can be shown[5] that, to a good approximation,

$$I(z)=I(0)\frac{\sin k(h-z)}{\sin kh}kh<3\pi/4,$$

and hence,

$$\frac{\int_0^h I(z)dz}{\int_0^h \cos kzI(z)dz}=\frac{\int_0^h \sin k(h-z)dz}{\int_0^h \cos kz\sin k(h-z)dz}$$

$$=\frac{1-\cos kh}{kh/2\sin kh}\quad kh<\frac{3\pi}{4}.$$

For greater lengths of antennas, it would be necessary to use more accurate current distributions[6] and evaluate these integrals numerically. Summarizing the results,

(a) $$\theta_n=\left\{\begin{matrix}0.59\frac{\lambda}{d} & n=1\text{ radians}\\ (n+\frac{3}{4})\frac{\lambda}{d} & n=2,3,4\cdots\text{ radians}\end{matrix}\right.$$

(b) $$\Delta\theta=\frac{\lambda}{d}\text{ radians}$$

(c) $$N=1.175\left(\frac{d}{\lambda}\right)^{\frac{1}{2}}\left|\frac{\int_0^h I(z)dz}{\int_0^h \cos kzI(z)dz}\right|^{\frac{1}{2}}-1$$

$$=1.87\left(\frac{d}{\lambda}\right)^{\frac{1}{2}}\left[\frac{1-\cos kh}{kh\sin kh}\right]^{\frac{1}{2}}-1\quad kh<3\pi/4 \tag{12}$$

(d) $$\theta_{\max}=1.175\left(\frac{\lambda}{d}\right)^{\frac{1}{2}}\left|\frac{\int_0^h I(z)dz}{\int_0^h \cos kzI(z)dz}\right|^{\frac{1}{2}}\text{ radians}$$

$$=1.87\left(\frac{\lambda}{d}\right)^{\frac{1}{2}}\left[\frac{1-\cos kh}{kh\sin kh}\right]^{\frac{1}{2}}\text{ radians}\quad kh<3\pi/4.$$

[5] D. Middleton and R. King, Quart. Appl. Math. 3, 302 (1946).
[6] R. King, *Graphical Representation of the Characteristics of Cylindrical Antennas*, Technical Report No. 20 (Cruft Laboratory, Harvard University, October, 1947).

IV. COMPARISON WITH EXPERIMENT

Figure 2 is a pattern measured by Meier and Summers[1] using an antenna of length $kh=2\pi$ (0.224) and ground screen diameter $d=6\lambda$. For this diameter $kd=2\pi(6)\cong 37$, and the inequality $kd\gg 1$ is not satisfied. Moreover, it is apparent that complete angular symmetry has not been preserved in the experimental work. However, there should be reasonably good agreement between this measured pattern and the theoretical predictions of Eq. (12).

Table I contains values computed from (12) and the corresponding quantities obtained from the measured pattern of Fig. 2.

Considering the fact that the inequality $kd\gg 1$ is not satisfied, the author feels that the agreement between the above values is quite satisfactory. The large discrepancy occurring in the values for θ_3 and θ_{max} is primarily due to the fact that the theoretical formulas are for small angles only, and the angles of greater than 30 degrees cannot be considered small. Of course, when the inequality $kd\gg 1$ is satisfied, the lobe structure will occur within small angles.

It was pointed out previously that the minimum points between the lobes should occur very close to the pattern for an infinite screen. It is worth while noting that this is the case in Fig. 2, at least for small angles.

V. CONCLUSION

Since most antenna measurements are made over large ground screens, the results derived in this and the author's previous paper are of considerable interest. Perhaps the most important information obtained is the rate of decrease of the finite screen effects as the diameter of the screen is increased. In particular, it was seen from these papers that

$$Z-Z_0\sim 1/d,$$
$$\theta_{max}\sim(1/d)^{\frac{1}{2}},$$

where $Z-Z_0$ represents the difference between the input impedance of the antenna over a finite and infinite screen.

It is apparent that these effects drop off very slowly and a very large screen is necessary to approximate an infinite screen. Of course, it is a well-known fact that placing an antenna off-center in a rectangular screen greatly reduces the changes of impedance. This technique, however, does not materially lessen the angle within which the lobe structure occurs and the formulas (12) give reasonably good estimates for screens of any shape when d is taken as a typical width.

To illustrate the very slow rate of decrease suppose a quarter-wavelength antenna were erected over a circular screen and it was desired that the lobe structure be contained within an angle of 25 degrees, i.e., $\theta_{max}=25°$. Then directly from the formula for θ_{max}, Eq. (12), it can be shown that the ground screen must be at least 32 wavelengths in diameter.

TABLE I.

	Theoretical	Experimental (left)	Experimental (right)
θ_1	5.6°	4.5°	6.0°
θ_2	16.7°	13.0°	15.0°
θ_3	26.3°	33.5°	37.5°
θ_{max}	42.0°	50.0°	55.0°
$\Delta\theta$	9.6°		
N	3.3	3	3

ACKNOWLEDGMENT

The author is indebted to Dr. R. W. P. King for the suggestion of this problem and for his many helpful discussions on the research associated with this paper.

Drift Velocity of Electrons in Argon and Argon Mixtures

J. M. KIRSHNER AND D. S. TOFFOLO
Naval Research Laboratory, Washington, D. C.

(Received February 13, 1952)

Since argon is prepared by fractional distillation from liquid air, the greatest impurity found in it is nitrogen whose boiling point lies close to that of argon. For this reason it is believed that some of the previous measurements on supposedly pure argon were actually made on nitrogen contaminated argon. By adding controlled amounts of nitrogen to pure argon, the drift velocity curves of various authors can be duplicated.

INTRODUCTION

DRIFT velocity measurements on argon have been made recently by Nielsen[1] in 1936, Herreng[2] in 1943, Hudson[3] in 1946, and Klema and Allen[4] in 1950.

The measurements were obtained by different methods. Nielson and Herreng obtained results in good agreement with each other, and Harriet Allen[5] using the theory of Morse, Allis, and Lamar[6] obtained a theoretical drift velocity curve for argon in very good agreement with the experimental results of Nielsen and of Herreng. Unfortunately neither of these authors explained their purification procedure; hence, some doubt was thrown on their results when Hudson, purifying his gas over calcium, obtained different results. Still later Klema and Allen, purifying their gas in the same manner as Hudson, made a number of self-consistent runs on argon at various pressures. Their curves are different from all the others. They also obtained drift velocity curves for a mixture of one percent nitrogen in argon. The curve for this mixture is in very good agreement with Hudson's curve for "pure" argon. Since argon is prepared by fractional distillation of liquid air, the greatest impurity found in it is nitrogen whose boiling point lies close to that of argon. The results reported herein indicate that Klema and Allen's gas may have had approximately 0.8 percent nitrogen in it and that when they added one percent of nitrogen, they then actually may have had a total of 1.8 percent nitrogen in their gas, which the authors believe was about the amount of impurity in Hudson's gas.

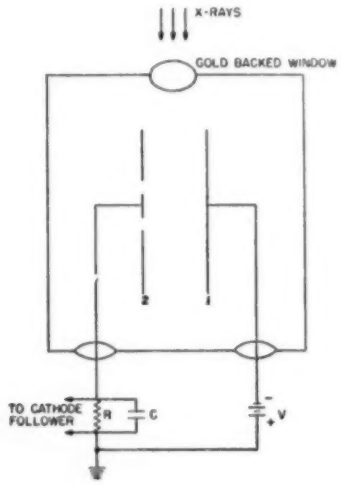


FIG. 1. Schematic of parallel plates and input circuit.

METHOD OF MEASUREMENT

The experimental method we used is essentially that of Hudson.[3] The gas in the volume between two parallel plates is ionized by the Compton electrons generated by a square wave of x-rays. Figure 1 shows a schematic diagram of the chamber used.

The window is made of glass backed with a thin gold film. Compton electrons coming mainly from the gold film ionize the gas between the plates.*

Because of the electric field between the plates the electrons move toward the collector electrode. The time of measurement is sufficiently short so that during the measurement the positive ions move only negligibly; consequently, the current in the external circuit depends only upon the motion of the electrons.

The signal from the parallel plate chamber is picked up by a cathode follower attached to the chamber, from which it goes into an amplifier having an amplification factor of about 10,000 and a signal delay of 0.145 microsecond. The signal then feeds into an oscilloscope which further amplifies and displays it. The rise time of the amplifier system is about 0.05 microsecond; the decay of a square wave input amounts to 1.2 percent of the initial height per microsecond.

The x-rays are produced by a pulse consisting of a 90 v square wave 200 μsec long having a rise time of 0.06 μsec which is fed to the grid of the x-ray tube. This pulse is fed simultaneously through a small condensor to the input of the oscilloscope where it forms a pip which together with the measured delay time of the amplifier establishes zero time for the output curve. The pulse is triggered at a 60 cycle rate.

[1] R. A. Nielsen, Phys. Rev. **50**, 950 (1936).
[2] Pierre Herreng, Compt. rend. **217**, 75 (1943).
[3] Donald E. Hudson, AEC Report, DDC 524.
[4] E. D. Klema and J. S. Allen, Phys. Rev. **77**, 661 (1950).
[5] Harriet Allen, Phys. Rev. **52**, 707 (1937).
[6] Morse, Allis, and Lamar, Phys. Rev. **48**, 412 (1935).

* The x-rays were introduced parallel to the plates since it is believed that this leads to a more uniform distribution of ions in the gas between the plates than can be obtained by allowing the x-rays to come from any other direction. For this method of ionization, little or no difference could be detected in the results when the position of the plates was interchanged relative to the x-ray beam (180° revolution).

The rise time of the x-rays as measured with an anthracene scintillation counter is about 0.1 μsec. Time markers are supplied by a transmitter at a frequency of 4.83 mc$\pm$15 kc/sec.

The input capacity to the cathode follower was measured with a Boonton meter and the input resistance with a Wheatstone bridge giving a time constant of 0.165 μsec.

EXPERIMENTAL PROCEDURE

A diagram of the chamber is shown in Fig. 2. The plates are of stainless steel, the spacers of porcelain, and the housing of brass. Electrical leads were introduced into the chamber through glass seals. The collector has a diameter of $\frac{3}{4}$ inch and the guard ring a diameter of three inches. The distance between the plates is $\frac{5}{8}$ inch. For this geometry the field between the collector and opposite plate is uniform to within one-tenth of one percent.[7]

In order to eliminate leakage and impurities caused by rubber gaskets, all seals in the entire system were soldered.

The chamber was thoroughly leak checked by a helium leak detector and soldered to the metal evacuating system which was evacuated to 0.01 micron while being baked out at 110°C.

Simultaneously a steel chamber containing calcium which was part of the system was heated to about 600°C.

After the system had been baked out and evacuated, argon, stated by the manufacturer to have less than 0.1 percent impurities, was allowed to enter the system, the temperature of the calcium first being allowed to drop to about 500°C. By cooling part of the system the gas was caused to convect over the hot calcium. The gas was allowed to convect overnight and the parallel plate chamber then detached and placed into the measuring system.

The temperature of the calcium was held at 500°C since nitrogen reacts actively with pure calcium only above about 410°C, although impurities in the calcium lower the necessary temperature. In particular the presence of calcium nitride acts as a catalyst so that after the reaction has started, absorption of nitrogen may be obtained at much lower temperatures; hence, one must be extremely careful in convecting over calcium, mixtures containing nitrogen. It was found that after some time in operation, observable absorption of nitrogen occurred at temperatures as low as 130°C.

For measurement of mixtures, the nitrogen was passed over calcium heated to 160°C and the temperature allowed to drop slowly until no further absorption with time was observed. This was then mixed with argon previously convected at 500°C. In order to add

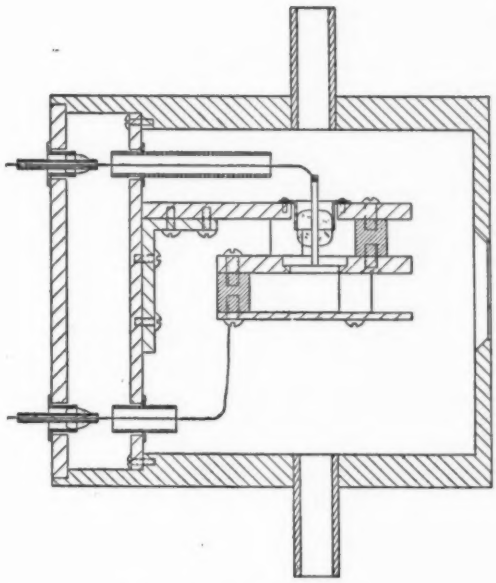


FIG. 2. Parallel plate chamber.

controlled amounts of nitrogen, an expansion chamber was used; the ratio of the volume of this chamber to the total volume of the system was measured by using a Wallace and Tiernan pressure gauge accurate to about 1 mm Hg. In order to get any small amount of nitrogen it is necessary to have the nitrogen in the total system at some pressure AR^n where A is the desired pressure and R is the ratio above. The expansion chamber is then closed and the rest of the system evacuated. The expansion chamber is then opened so that the pressure becomes AR^{n-1}. Continuation of this process for n expansions eventually gives A, the pressure desired.

Although the chamber was detached from the purification system to make measurements, this procedure is justified in that measurements taken immediately after detachment and several days later showed no appreciable difference.

THEORY

Hudson[3] solved the equations for the response of the above type of parallel plate chamber to a square wave. We have extended this solution to the case of a general source function of time which is zero for times less than zero. In particular, these equations have been solved for the case of a source function with an exponential rise.

Consider Fig. 1 which shows the parallel plates in series with the input to the cathode follower.

A single electron moving in the field between the two parallel plates has work done on it

$$VI(t)dt = Eedx,$$

where $E = V/d$ is the field strength if the drop across R is sufficiently small.

[7] E. B. Rosa and N. E. Dorsey. Bull. Natl. Bur. Standards, 3, 534 (1907).

Hence, if d is the distance between the plates,

$$I(t)\dagger = we/d, \tag{1}$$

where $w = dx/dt$ = constant is the drift velocity of an electron.

Given a density distribution $\rho(x, t)$ of electrons between the plates, the current $I(t)$ is then

$$I(t) = \frac{we}{d}\int_0^d \rho(x, t)dx, \tag{2}$$

where x is the perpendicular distance of the electron from plate 1.

Now $\rho(x, t)$ is given by the equation of continuity,

$$\frac{\partial \rho}{\partial t} + w\frac{\partial \rho}{\partial x} = f(t), \tag{3}$$

the solution of which is

$$\rho_1(x, t) = \int_0^t f(t-\tau)d\tau; \quad t \leq x/w, \tag{4}$$

$$\rho_2(x, t) = \int_0^{x/w} f(t-\tau)d\tau; \quad t \geq x/w. \tag{5}$$

The current $I(t)$ is then given by,

$$I_1(t) = \frac{we}{d}\left\{\int_0^{wt} \rho_2 dx + \int_{wt}^d \rho_1 dx\right\}; \quad t \leq \frac{d}{w}, \tag{6}$$

$$I_2(t) = \frac{we}{d}\int_0^d \rho_2 dx; \quad t \geq \frac{d}{w}. \tag{7}$$

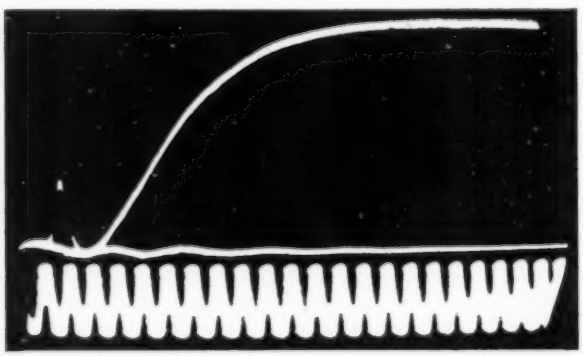


FIG. 3. Typical voltage-time curve for E/p less than 2. The small pip is the zero time indicator.

† Actually, in order to obtain a uniform field a guard electrode is used. In this case the electron induces current on the guard ring as well as on the collector and the fraction of current induced on the collector to that induced on the guard ring depends on the distance of the electron from the collector; however, if the gas between the guard electrode and the plate is ionized uniformly as well as that between the collector electrode and the plate, this effect is compensated to a large extent.

The voltage appearing across the resistor R is given by the circuit equation

$$I(t) = C\frac{dV}{dt} + \frac{V}{R}; \quad T \equiv RC, \tag{8}$$

whence,

$$V_1(t) = \frac{e^{-t/T}}{C}\int_0^{d/w} e^{\tau/T} I_1(\tau)d\tau; \quad t \leq \frac{d}{w} \tag{9}$$

$$V_2(t) = \frac{e^{-t/T}}{C}\int_{d/w}^t e^{\tau/T} I_2(\tau)d\tau + V_1\left(\frac{d}{w}\right)e^{d/wT - t/T}; \quad t \geq \frac{d}{w}. \tag{10}$$

Now if $\phi(t)$ is the response function of the amplifier to a unit square wave, then the output of the amplifier $\psi(t)$ (the signal observed on the oscilloscope) is given by

$$\psi(t) = \int_0^t \dot{V}(\tau)\phi(t-\tau)d\tau. \tag{11}$$

In particular if the function $\phi(t)$ is an exponential decay as is usually true and if the rise time of the amplifier is sufficiently small compared to the input rise time of the cathode follower, then

$$\phi(t) = Ae^{-t/\beta},$$

where A is the amplification factor of the amplifier and β is the decay constant.

The rise time of the amplifier used was about $\frac{1}{3}$ of that of the cathode follower.

Since the effect of two rise times in cascade is approximately equal to the square root of the sum of the squares of the individual rise times, neglect of the smaller rise time introduces an error of about ten percent in the correction terms. Since, however, these terms are less than five percent of the original terms, the total error caused by neglect of this rise time is about $\frac{1}{2}$ of one percent.

For

$$\phi(t) = Ae^{-t/\beta}$$

we can integrate (11) by parts to obtain

$$\psi_1(t) = AV_1(t) - \frac{A}{\beta}\int_0^t V_1(\tau)e^{-(t-\tau)/\beta}d\tau; \quad t \leq \frac{d}{w}, \tag{12}$$

$$\psi_2(t) = AV_2(t) - \frac{A}{\beta}\int_0^{d/w} V_1(\tau)e^{-(t-\tau)/\beta}d\tau - \frac{A}{\beta}\int_{d/w}^t V_2(\tau)e^{-(t-\tau)/\beta}d\tau; \quad t \geq \frac{d}{w}. \tag{13}$$

This is the complete solution for an arbitrary function of time which is zero for times less than zero.

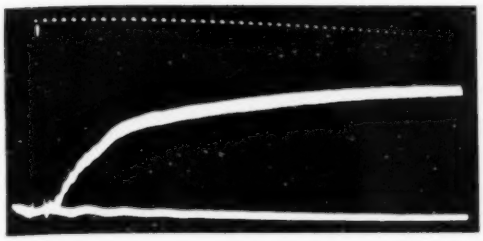


FIG. 4. Excitation of metastable states in argon at $E/p=3.4$.

For the particular case of

$$f(t)=a(1-\epsilon^{-t/\sigma}); \quad t\geq 0, \tag{14}$$

where σ is the rise time of the x-rays and a is the rate of production of ions after a sufficiently long time:

$$\psi_1(t)=\frac{Awea\beta T^2}{Cd(\sigma-T)}\left\{\frac{d+wT}{T-\beta}[\beta+T(\epsilon^{-t/T}-1)]\right.$$
$$+\frac{\sigma(d+w\sigma)}{T(\sigma-\beta)}[\sigma(1-\epsilon^{-t/\sigma})-\beta]+\frac{(\sigma-T)}{T}$$
$$\left.\times\left[w(\beta-t)-\frac{\beta^2(d+w\beta)}{(\sigma-\beta)(T-\beta)}\epsilon^{-t/\beta}\right]\right\}; \quad t\leq\frac{d}{w} \tag{15}$$

$$\psi_2(t)=\frac{AweaT^2}{Cd(\sigma-T)}$$
$$\times\left\{\frac{w\beta}{T(T-\beta)}\left[(\sigma^2-T^2)+\frac{\sigma_2}{T}(\sigma-T)\right]\epsilon^{(d/wT)-(t/T)}\right.$$
$$+\frac{\beta^2(d+w\beta)(T-\sigma)}{T^2(T-\beta)(\sigma-\beta)}\epsilon^{-t/\beta}+\left[(d+wT)\left(\frac{d}{wT}-\frac{\beta}{T}-1\right)\right.$$
$$+\frac{\sigma}{T^2}(d+w\sigma)\left(\sigma-\frac{d}{w}+\beta\right)+(\sigma-T)$$
$$\left.\left.\times\left(\frac{d^2-\beta wd+\beta^2w^2-\sigma^2w^2}{wT^2}-\frac{w\sigma+wT}{T-\beta}\right)\right]\epsilon^{(d/w\beta)-(t/\beta)}\right\}$$
$$t\geq d/w.$$

Where to simplify to some extent the rather tedious algebra, the following approximations have been made. Since $\epsilon^{-t/\sigma}$, $\epsilon^{-t/T}$ for our system $\ll 1$ for $t\geq d/w$,

$$I_2(t)=\frac{wea}{d}\left\{\sigma\epsilon^{-t/\sigma}(d+w\sigma)-\sigma^2w+\frac{d^2}{2w}\right\}\cong\frac{wea}{d}\left\{\frac{d^2}{2w}\right\}$$

$$V_1\left(\frac{d}{w}\right)\cong\frac{eaT}{2Cd}[d^2-2w^2(\sigma^2+\sigma T+T^2)],$$

and terms of $\psi_2(t)$ containing the exponentials

$$\exp\left(-\frac{t}{\beta}+\frac{d}{w\beta}-\frac{d}{wT}\right) \quad \text{and} \quad \left(-\frac{d}{w\sigma}-\frac{t}{\beta}+\frac{d}{w\beta}\right)$$

have been neglected.

From the equations it can be shown that the current in the case of a square wave source function must reach a steady state at the time it takes an electron to cross the distance between the plates. To the order of approximation used, a steady state is reached at this time also for our case of a finite but small rise time. One can observe further that if the attachment coefficient remains constant in those gases in which attachment takes place, the steady state must still occur at this same time.

If the time constants of the circuit are sufficiently small, the general shape of the output signal will follow closely that of the current, however, the output signal will not reach its maximum until some time later than the time of transit of an electron so that it is necessary to analyze the shape of the output signal to determine the transit time.

Method of Calculation

$\psi(t)$ can be written in terms of the variable t and the unknown w as $K\bar{\psi}(w, t)$ where K is a constant. We evaluated w by a method due to Hudson.[3]

To do this we find the maximum value of $\bar{\psi}(w, t)$. Call this $\bar{\psi}(w, t_m)$. We also find $\bar{\psi}(w, d/w)$ and form the ratio

$$R(w)=\frac{\bar{\psi}(w, d/w)}{\bar{\psi}(w, t_m)}.$$

Using values of w which cover the range of possible values we evaluate a set of $R(w)$. We set up a table of w and $R(w)$. Once this has been done, the drift velocity w can be found quite rapidly as follows.

Assume some ratio $R(w_0)$. One finds the point on the curve having a height of $R(w_0)$ times the maximum height. From this one finds the time t_1 corresponding to $R(w_0)$; but $w_1=d/t_1$.

One now looks on the table for $R(w_1)$, finds t_2 from the curve and continues this process until $R(w_n)=R(W_{n+1})$. In general the third approximation is sufficient.

Results and Discussion

For pure argon at an E/P less than 2, and for nitrogen contaminated argon, oscilloscope traces like that of Fig. 3 were obtained. For an $E/P>2$, traces like that of Fig. 4 were obtained for pure argon. The presence of very small amounts of oxygen (order of 0.01 percent)

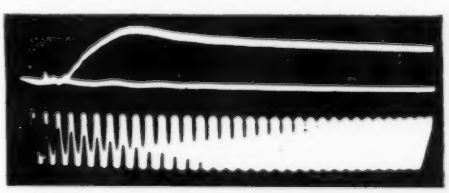


FIG. 5. Effect of small amounts of oxygen and consequent attachment on the voltage-time curve.

gave a trace like that of Fig. 5. The presence of 0.1 percent oxygen in our gas completely eradicated the signal as far as our amplification system was concerned.

Our pure argon curve, Fig. 6, is in very good agreement with the experimental curves of Herreng[2] and Nielson[1] and with the theoretical calculations of Harriet Allen[5] based on the theory of Morse, Allis, and Lamar.[6]

Our curve of argon containing 0.7 percent nitrogen (Fig. 7) is in good agreement with the "pure" argon curve of Klema and Allen.[4]

The addition of 1.8 percent nitrogen (Fig. 7) gives a curve in good agreement with Hudson's curve[3] for "pure" argon and with Klema and Allen's curve[4] for a one percent nitrogen in argon mixture.

As previously discussed under theory, if the attachment coefficient is constant, the general shape of the curve should be the same as for no attachment, it is therefore clear that in Fig. 5 the attachment coefficient

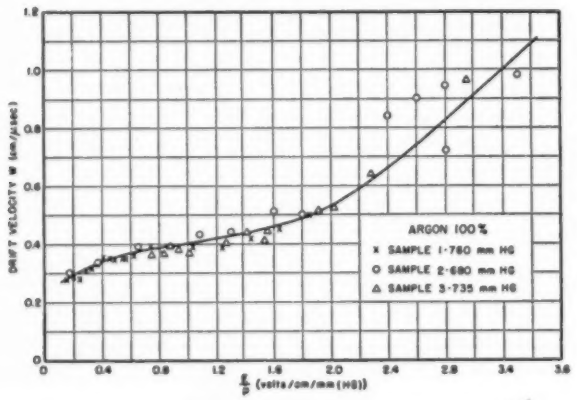


FIG. 6. Drift velocity of electrons in pure argon as a function of E/p.

is changing with time and furthermore that it is increasing.

Since excited oxygen molecules and oxygen atoms are better attachers than normal oxygen molecules[8,9] the increase of attachment with time may be the result of the excitation and dissociation of oxygen molecules by x-rays and Compton electrons, equilibrium finally being reached when the number of molecules being excited and dissociated is equal to the number dropping back to the ground state and recombining.

The trace of Fig. 4 for argon, where $E/P>2$ can be explained as due to excitation of metastable states. In

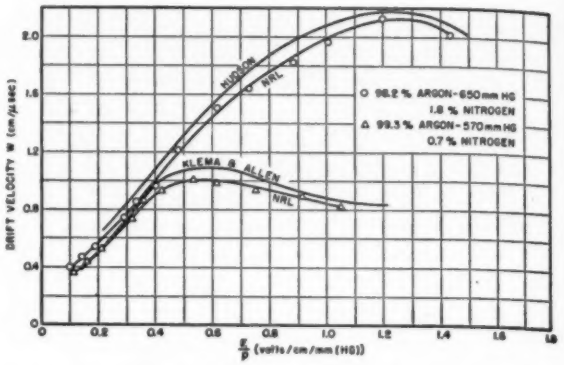


FIG. 7. Comparison of NRL argon plus nitrogen curves with "pure" argon curves of Hudson and of Klema and Allen.

this case the "knee" of Fig. 5 corresponds to the beginning of the steady state for Fig. 3. Electrons moving through the gas in this case have enough energy to excite some of the atoms to metastable states. Hence, at the time when the steady state would normally begin, there are present in the chamber a number of excited atoms. Passage of further electrons through the chamber ionizes some of these excited atoms so that the total number of electrons increases.

The presence of these metastables is an indication of the purity of the gas since the presence of any impurities with low levels would absorb the electron energy and prevent excitation of the metastable states of argon. Specifically, the addition of 0.1 percent N_2 completely destroys this phenomenon.

When E/P, for argon, is greater than 2, our equations no longer fit the curve and our method of finding the drift velocity is no longer valid. In these cases, therefore, we took the time at the beginning of the "knee" as the time of transit of the electrons across the chamber. This, of course, means that measurements in this region are not very accurate. Furthermore, as E/P increases, the "knee" becomes more difficult to identify so that our values decrease in accuracy.

However, in those cases where our equations hold, the time at which steady state occurs on the output curve differs by less than 10 percent from the transit time. Consequently, use of the "knee" probably causes an error of less than ten percent where this "knee" is well defined and perhaps an error of twenty percent for the region of E/P around 3. We believe that our measurements of drift velocity for E/P less than 2 are accurate to within 5 percent.

We would like to express our appreciation to Mr. Sydney Rattner who designed much of the required electronic equipment.

[8] N. E. Bradbury, Phys. Rev. 44, 883 (1933).
[9] L. B. Loeb, *Fundamental Processes of Electrical Discharge in Gases* (John Wiley and Sons, Inc., New York, 1937), p. 301.

Letters to the Editor

Thermionic Emission and Electrical Conductivity of Oxide-Coated Cathodes

SHIN-ICHIRO NARITA
Kobe Kogyo Corporation, Kobe, Japan
(Received January 14, 1952)

THERMIONIC emission and electrical conductivity of "sintered" oxide-coated cathodes have been measured by means of an imbeded probe and comparison made with that of ordinary oxide-coated cathodes. The sintered cathode was prepared by converting barium carbonate to the oxide under conditions of slow speed of exhaust and at very high cathode temperature.

The activities of the sintered and ordinary cathodes were varied by evaporating barium from a batalum getter onto the surfaces.

Figure 1 shows the temperature dependence of thermionic emission and electrical conductivity of both sintered and ordinary oxide cathodes in the initial and the fully activated state (Table I). Their emissions and conductivities at 785° K, work functions

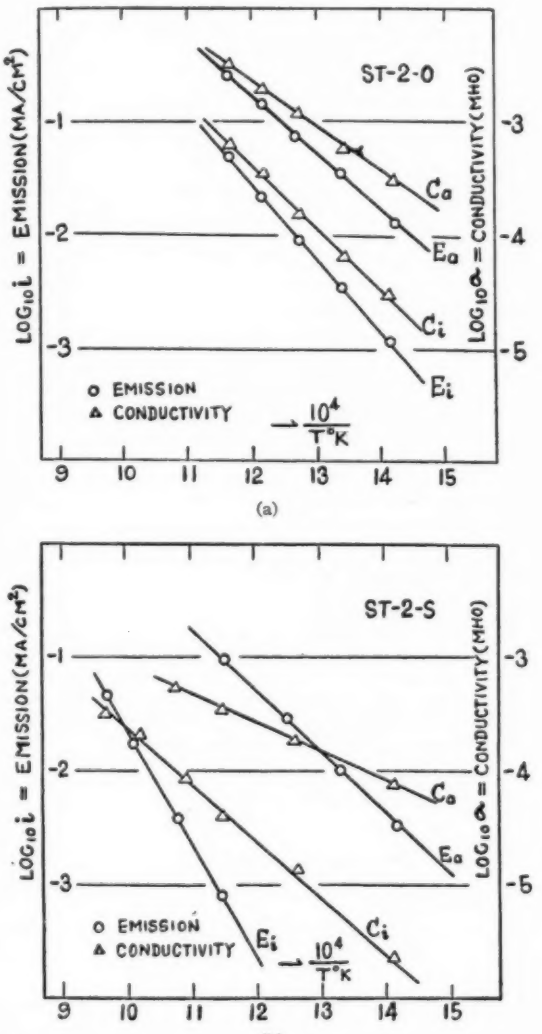


FIG. 1. Logarithmic emission and conductivity *vs* $1/T$ plots of ordinary (a) and sintered (b) cathode. (a) E_i, C_i: emission and conductivity at the initial state. (b) E_a, C_a: emission and conductivity at the activated state.

TABLE I.

Cathode		i mA/cm²	σ Mho	Φ ev	E ev	$\tilde{\chi}$ ev
Ordinary	Initial	9×10^{-3}	1.5×10^{-4}	1.20	1.04	0.16
	Activated	8.1×10^{-1}	1.2×10^{-3}	1.03	0.86	0.17
Sintered	Initial	7.0×10^{-5}	1.2×10^{-6}	1.89	0.94	0.95
	Activated	2.8×10^{-1}	1.7×10^{-4}	1.06	0.44	0.62

Φ, activation energies of conductivity E, and the apparent values of electron affinity $\tilde{\chi}$ obtained from the differences between Φ and E are tabulated.

It is to be noted that $\tilde{\chi}$ of the sintered cathode is very large compared with that of the ordinary cathode, especially at the initial state, and moreover it varies with the state of cathode activity.

The former may be explained by the difference in structure between the sintered cathode and the ordinary one. The oxide particles in ordinary cathodes are considered to have distinct grain-to-grain boundaries. Therefore the contact resistance between grains is so high that the measured activation energy of conductivity is either that of electron conduction through pores in the coating as advocated by Loosjes and Vink[1] or is the height of contact potential barrier between grains. These values differ from the activation energy of conduction through the oxide grains which perhaps is measured in the case of sintered cathodes.

The change of $\tilde{\chi}$ with the state of cathode activity may result from the formation of a barium later on the oxide surface. However, no change of $\tilde{\chi}$ has been observed by others;[2] this has been one of the powerful proofs to deny the effect of the surface barium assumed in early days by Becker[3] and others. This inconsistency may come from the fact that values of $\tilde{\chi}$ have usually been obtained from measurements with ordinary cathodes.

Also in this experiment, when barium was evaporated onto the cathode surfaces at very low temperature (<550°K), a small activation energy of conductivity (less than 0.1 ev and similar to that Vink-Loosjes reported[1]) was obtained from both sintered and ordinary cathodes. This conduction was easily diminished and caused to vanish by heating the cathode. This was accompanied by barium evaporation from the cathodes which was detected by measuring the increase of thermionic emission from a tungsten wire located near the cathode. Such small values of the activation energy for conduction through the grains is difficult to understand, and this conduction is believed to be a surface conduction through a multiatomic barium layer on the surface of the oxide particles. This has been confirmed by comparing it with the conduction through a very thin barium film deposited on a silica plate.

From these results it is believed that in the oxide-coating there may be three mechanisms of electron conduction: (a) conduction by an electron gas in the pores or conduction through grains limited by contact potential barriers between grains, (b) pure conduction through grains, and (c) surface conduction through surface barium layers. All of these are connected in parallel in the most general case, and the monoatomic surface layer of barium serves to reduce the work function.

[1] R. Loosjes and H. J. Vink, Philips Research Repts. **4**, 449 (1949).
[2] E. Nishibori and H. Kawamura, Proc. Phys-Math. Soc. Japan **22**, 378 (1940).
[3] J. A. Becker and R. W. Sears, Phys. Rev. **38**, 2193 (1931).

Errata: General Theory of Electromagnetic Horns

[J. App. Phys. **22**, 1447 (1951)]

A. F. STEVENSON
Department of Mathematics, Bishop's University, Lennoxville, Quebec, Canada

PAGE 1449, Eq. (2.16): for H_r read E_r. Page 1457, 2 lines below Eq. (6.6): for a read z. Page 1457, last two equations: interchange n and m and replace a by b. Also in the last equation on this page a minus sign should be inserted before the expression

on the right-hand side. Page 1458, Eq. (6.7): interchange n and m and replace a by b. Also in the equation for h in the first solution and in the equation for e in the second solution, the signs of the expressions on the right-hand sides should be changed. Page 1458: in the expression which occurs 4 lines above Eq. (6.7), n and m should be interchanged and a replaced by b. Page 1458, fifth line from bottom of first column: for $ka>n\pi$, read $kb>m\pi$.

Erratum: The Thermal Conductivity of Metals at High Temperature

[J. Appl. Phys. 23, 177–180 (1952)]

C. L. Hogan and R. B. Sawyer
Lehigh University, Bethlehem, Pennsylvania

THE figures originally submitted with the article were unsuitable for publication, and new ones had to be drawn. Unfortunately, neither Mr. Hogan nor Mr. Sawyer noticed that on the new figures the captions for Fig. 6 and Fig. 7 were interchanged. The correct captions for these figures are as follows: Fig. 6. Lorentz plot (k *vs* σT) for stainless steel type 466. Fig. 7. Lorentz plot (k *vs* σT) for Inconel "*X*."

Erratum: Review of Book Entitled "Microwave Electronics"

[J. Appl. Phys. 23, 293 (1952)]

F. Carl Willenbrock
Harvard University, Cambridge, Massachusetts

THE name of the McGraw-Hill Book Company was given in error as publishers of the book by John C. Slater, *Microwave Electronics*. This book was published by D. Van Nostrand Company, Inc., New York City.

The Journal regrets any inconvenience this error may have caused.

The Application of Nutting's Equation to the Viscoelastic Behavior of Certain Polymeric Systems

G. W. Scott Blair
National Institute for Research in Dairying, University of Reading, England
(Received January 21, 1952)

UNDER this title Buchdahl and Nielsen[1] have discussed the proportionality between the creep function and the parameter defined by them as the "dissipation factor." The latter is defined as E''/E', where E'' and E' are the real and imaginary parts of a complex modulus. The former is obtained from the frequency of vibration, and the dissipation factor is obtained from the decrement of vibration.

From an equation of Zener's the authors derive

$$n \doteq 2E''/\pi E', \qquad (8)^1$$

where n is the slope of log strain/log time curve for constant stress, previously defined (with the symbol k) by Scott Blair and Coppen[2] as the "dissipation coefficient." Buchdahl and Nielsen point out that "the dissipation factor measures the ratio of the dissipation of energy to the total stored energy when a solid is subject to stress (or strain)."

The distinction between "dissipation factor" and "dissipation coefficient" is important, but even if this is appreciated, it would appear from the sentence quoted, taken in conjuction with Eq. (8),[1] that the dissipation *coefficient* should also measure the ratio dissipated to total energy.

Although in the first paper by Scott Blair and Coppen[2] this was implied to be the case, it was later realized that the inference is in general incorrect. Thus Scott Blair and Veinoglou[3] describe the behavior of pastes "which behave *under compression* like an elastic solid, i.e., they are strained to an extent proportional to the applied stress and the deformation is almost instantaneously completed" (i.e., the dissipation coefficient does not differ significantly from zero). "They do not, however, recover when the stress is removed."

Since such meterails exist and since there are doubtless many others for which the discrepancy, though less pronounced, is of importance, it is felt that Eq. (8)[1] should be accepted with caution. Its validity depends on the exact significance of the symbol "$\doteq$." If this is taken to mean "in the limiting case *of certain materials*," the equation is, no doubt, correct. But the authors' statement that, "there can be no doubt about about the basic interdependence of these quantities [E''/E' and n]," cannot be accepted as a generalization.

I am indebted to Dr. Buchdahl for his courtesy in discussing this matter with me by private correspondence.

[1] R. Buchdahl and L. E. Nielsen, J. Appl. Phys. 22, 1344 (1951).
[2] G. W. Scott Blair and F. M. V. Coppen, Proc. Roy. Soc. (London) 128B, 109 (1939).
[3] G. W. Scott Blair and B. C. Veinoglou, J. Sci. Instr. 21, 149 (1944).

The Application of Nutting's Equation to the Viscoelastic Behavior of Certain Polymeric Systems

Rolf Buchdahl and Lawrence E. Nielsen
Monsanto Chemical Company, Plastics Division, Springfield, Massachusetts
(Received February 18, 1952)

WE agree with Professor G. W. Scott Blair that the relationship between mechanical damping and creep rate[1] is not universally true, and we stated so in our paper. However, we feel in many cases, even for widely different types of materials, that it is at least approximately correct. Besides the examples listed in our paper, there are other cases where a similar relationship between the dissipation factor (mechanical damping) and a parameter describing the time dependence of creep or stress relaxation has been found. Experimental evidence of such relationships have been found for glass,[2] metals,[3,4] and high polymers such as rubbers.[5] It seems to us that a relationship covering such a wide range of materials must have some general validity, although admittedly, it might not include everything.

On the theoretical side the assumptions used in deriving Zener's equation (on which our equation is based) are quite general in nature.

[1] R. Buchdahl and L. E. Nielsen, J. Appl. Phys. 22, 1344 (1951).
[2] W. A. Weyl, Research 1, 50 (1947).
[3] T. S. Kê, Phys. Rev. 71, 533 (1947).
[4] A. S. Nowick, J. Appl. Phys. 22, 925 (1951).
[5] Tobolsky, Dunell, and Andrews, Textile Research J. 21, 404 (1951).

X-Ray Line Broadening of Plastically Deformed Marble

D. Rosenthal and M. Kaufman*
University of California, Los Angeles, California
(Received January 21, 1952)

TWO major factors have been suggested to account for the x-ray line broadening caused by cold working of polycrystalline metals.

(1) The breaking up of the original crystals, or grains, into crystallites small enough to produce the effect of particle size broadening[1];

(2) The lattice distortion, i.e., the variation of the lattice parameter from grain to grain, or even within a single grain, due to residual microstresses.[2]

Experimentally, both grain diminution and lattice distortion have been observed during plastic deformation. The question

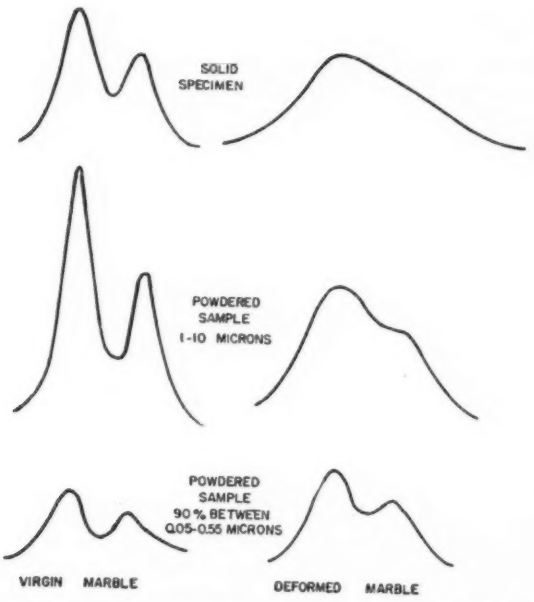


FIG. 1. Photometer traces of (651) planes of marble specimens. Cu radiation; one sheet of 0.00035-inch Ni foil over film. Various exposure times.

arises which of the two is the more important factor in line broadening. Recent experimental results point toward lattice distortion.[3] But the results are not direct and lack the evidence of what may be called a crucial test. To produce such a test it is necessary to relax the residual stress within the crystallites without affecting their size. If the line broadening decreases markedly after the relaxation, then the lattice distortion is the major factor; if the line broadening persists after the relaxation, then the particle size is the predominant effect. The difficulty of the test resides, of course, in the achievement of the required type of relaxation. Low temperature annealing, or recovery, while removing the lattice distortion, is not immune from recrystallization and grain growth on the microscale. On the other hand, mechanical separation of the crystallites cannot be achieved in metals without further plastic deformation. However, this last difficulty can be minimized, if not altogether eliminated, by using plastically deformed marble instead of metals.

It is known[4] that marble deforms plastically under uniaxial load if it is confined by a large hydrostatic pressure, of 1000 atmospheres and more, but that it fails brittly under the same type of load if the confining pressure is reduced to one atmosphere. Thus, it becomes possible to separate mechanically the crystallites of a plastically predeformed marble without further plastic deformation merely by crushing the sample in a mortar.

In the present test, samples of Yule marble were used† which have undergone a plastic deformation of from 15 percent to 20 percent in tension and compression under a confining hydrostatic pressure of 10,000 atmospheres. The stress-strain diagrams are very similar to those shown by metals. Likewise, the mode of deformation appears to be one of twin and translation, rather than twin alone.[5] X-ray diffraction diagrams of the deformed specimens reveal considerable line broadening as compared to the diagrams of the undeformed virgin marble, Fig. 1.

TABLE I. X-ray line broadening of marble specimens.

State of aggregation	rms angular line width (degrees) Virgin	Average	Deformed	Average
Solid specimen	0.24 0.26	0.250	0.57 0.62 0.63 0.67	0.622
Grain size: 1–10 microns	0.29 0.30	0.295	0.48 0.55 0.46 0.48 0.58 0.60 0.62 0.49	0.540
Grain size: 90% between 0.05–0.55 micron	0.39 0.28 0.32 0.51(?)	0.375	0.38 0.33 0.42 0.45 0.44 0.47	0.415

For the purpose of further comparison both deformed and virgin specimens were reduced by crushing to a powder. From this powder samples of various grain sizes for x-ray diffraction diagrams were obtained by means of levigation in water using Stokes' formula. A sample of a given grain size was obtained by repeated crushing of the precipitate having a larger grain size. In this way a representative sample of the original material was available for the x-ray diffraction. To minimize the influence of heat the process of desiccation was carried out at room temperature. All operations were performed in duplicate. The grain size of samples levigated below one micron was checked by the electron microscope.§

In Fig. 1 the line broadening of samples prepared from plastically deformed specimens has been compared to that of samples of the same grain size obtained from the virgin marble. It is seen that comparatively there is still considerable line broadening in the plastically deformed sample when the grain size averages more than one micron, but there is practically none when the grain size averages less than one micron. For a more quantitative comparison the results of several measurements of line broadening are expressed in Table I in terms of rms angular width following a procedure described elsewhere.[6]

It is thus evident that the x-ray line broadening of the plastically deformed marble is due very little to the particle size effect. The inference is that the observed line broadening is caused by microstresses which are being relaxed when the specimen is reduced below a certain critical size. This size may be considered as the dimension of the coherently scattering domain or crystallites into which the original grain has been subdivided as a result of plastic deformation. In the present case the size of the domain has been found to be around 3000A, a value which appears to be consistent with the Bragg's criterion for yielding based on dislocations.[7]

* Department of Engineering, University of California, Los Angeles. Work sponsored by the Institute of Geophysics, University of California.

† Courtesy of Dr. D. Griggs, Institute of Geophysics, University of California at Los Angeles.

[1] W. A. Wood and W. A. Rachinger, J. Inst. Metals London, **75**, 571 (1949).
[2] H. Lipson and A. R. Stokes, Nature **163** 871 (1949).
[3] B. E. Warren and B. L. Averbach, J. Appl. Phys. **21**, 595 (1950).
[4] Th. van Karman, Forschungsheft 118 Verein Deutscher Ingenieren, 1911.
[5] D. Griggs and W. B. Miller, Bulletin Geological Soc. **62**, 853 (1951).
[6] D. Rosenthal, J. Appl. Phys. **20**, 1257 (1949).
[7] Sir Lawrence Bragg, "Effects Associated with Stresses on a Microscopic Scale," Symposium on Internal Stresses, The Institute of Metals, 1948, p. 221.

§ Courtesy of H. Froula, Dept. of Engineering.

Books Reviewed

The Quantum Mechanics of Particles and Wave Fields

By Arthur March. Pp. 292+x, Figs. 5. John Wiley & Sons, Inc., New York, 1951. Price $5.50.

Dr. March has written an elementary textbook on quantum mechanics which is novel in the choice of topics and in the distribution of material among the topics. The first five chapters are devoted to the nonrelativistic quantum mechanics of particles. This treatment is quite brief, however, it covers some items which are often omitted from an elementary text. One such item, which is particularly well done, is transformation theory. On the other hand, perturbation theory, which is treated well in many texts, receives a very brief treatment. The remainder of the book is devoted to relativistic theory and to the quantum theory of fields. The treatment of these subjects is noteworthy for the inclusion of material on spin-zero and spin-one particles, and the application of this material to nuclear theory.

By means of this choice and distribution of material a gap in literature of quantum mechanics has been filled. It is now possible to gain a speaking acquaintance with the terminology and methods of the quantum theory of fields without going to the advanced texts. This is clearly of value to the experimentalist who must correlate this data with theory but in addition, the student who intends to specialize will find the climb to a working knowledge of field theory lightened by the addition of a step between the weak treatments found in most elementary texts and the abstract treatments found in the advanced works.

I cannot close without mentioning the treatment of the pseudoscalar field of which March is clearly and justifiably proud. The concept of a pseudoscalar is usually introduced as a one component quantity which transforms as a scalar under proper rotations but which reverses sign under reflections. The connection of these transformation properties with tensor analysis is nebulous and is made more so by denoting scalars and pseudoscalars by similar notations. This is avoided by introducing the pseudoscalar as an antisymmetric tensor of rank four and deducing its properties from this definition. March retains the notation of tensor analysis even in the application of the pseudoscalar field to meson theory, thus making clear the origin of the differences between the effects of the two types of meson theories. The pseudovector field is treated similarly though in less detail. This feature alone makes the book well worth reading for anyone who is unsure of the exact nature of the pseudoscalar field.

F. J. Milford
Massachusetts Institute of Technology

Finite Deformation of an Elastic Solid

By Francis D. Murnaghan. Pp. 140. John Wiley & Sons, Inc., New York, 1951. Price $4.00.

A unified treatment is given of the influence of squares and higher powers of the strain components in the theory of elasticity. An excellent collection of exercises or problems forms an essential part of the text. Thought provoking "why" or "show this" comments will help most students to understand the fundamentals.

Matrix notation and concepts are developed quite fully in the first chapter and used almost exclusively from then on. The only exception is in the treatment of the general equations of compatibility by means of tensors. A careful discussion of finite strain based on the initial reference frame (Lagrangian definition) includes the form and significance of the invariants. Stress is then defined in the Eulerian sense. The theorem of virtual work thus involves a term much more elaborate than the simple product of stress and strain which appears in the infinitesimal theory. Elastic strain energy, a function of the components of strain, is taken as energy of deformation per unit of initial volume and the stress strain relation is derived. The simple gradient relation of elementary theory no longer holds.

Elastic isotropy is examined at length and the various order approximations are described. Comparison of an integrated linear theory and of the second-order approximation with Bridgman's extreme pressure tests shows that for really precise agreement, third-order elastic constants would be needed. Nonisotropic elastic media of various symmetries are analyzed for various powers of strain in the energy function and relations obtained between the elastic constants.

Stress is then referred to the undeformed medium and simple shear and tension are studied in the presence of initial stress.

The book closes with a treatment of several problems: the compression of a spherical shell, the compression of a circular tube, and the torsion of a circular cylinder.

Surprisingly there is no reference to the extensive literature on the subject; the author is content to refer the reader to recent issues of reviewing journals. There seem to be only two questionable features of the development aside from the matter of the over-all convenience of the particular formulations employed. One is an insistence on consistency of order of approximation. It would appear completely permissible to choose a less complete set of second-order terms plus a third-order term if this happens to be convenient. The other disturbing point is the statement that macroscopic elastic concepts alone are able to explain the phenomenon of flow in metals. However, these are minor points in a well-organized and well-written book.

D. C. Drucker
Brown University

Helicopter Analysis

By Alexander A. Nikolsky. Pp. 340+ix. John Wiley & Sons, Inc., New York, 1951. Price $7.50.

Professor Nikolsky's book concerns the aerodynamical and mechanical analysis and design of helicopter rotors. He treats an extraordinarily intricate subject in a practical and straightforward way. The work appears designed both as a technical guide for the engineer's desk and as student's text. In the former capacity it is exemplary, setting forth the relevant practical problems and techniques for their solutions. This aspect of the book is beneficial in the latter capacity too, although gaps are left for the oral teacher's more perspicuous exposition.

The text covers forward and hovering flight, blade feathering and flapping, flight and stability analyses, and a chapter is devoted to the effects of blade flexibility. It embodies a well-selected resume of the practical literature as well as the author's able original contributions. To attain completeness, the opening chapter reviews classic airscrew theory and another treats of constant coefficient linear differential equations. The presentation is utile and down-to-earth; avoided is the obfuscation which besets so many otherwise decent and candid people when they write on this subject.

Helicopter theory is basically routine mechanics and aerodynamics rendered obtuse by, first, certain theoretic difficulties, and, second, the prolixity of factors which must be taken into account. Hence, the salient task of the investigator-expositor to assess and appraise these aspects. Should they be approximated or circumvented? By what means and by what justification?

The theoretic difficulties center, as often, in nonlinear differential equations. In helicopter work, they appear when we try to compute the induced air flow and in a form abstruse enough to sever any hope of progress. For both wings and propeller blades there exist elegant theories of optimal design (in the sense of minimal induced drag), but the analogous development for helicopter rotors appears immitigable. Professor Nikolsky opens his work with a cursory review of latter successful theory, but subsequently makes no use of it. He can't, of course. Still a qualitative discussion of the difficulties and a justification for their circumvention is possible. We feel something more than the

author's mere adumbration of this and other forced simplifications would be desirable.

The profusion of factors which enter the calculation is evidenced by the protracted tables of notation which are preludes to many chapters. (Happily all the entries are not always invoked later on.) Besides an abundance of physical variables proper, certain combinations are grouped together and granted names. (Solidity ratio, figure of merit, etc.) They facilitate comparison between designs and with test data. But at times they lend a deceptive complexity to the formal mathematics. The novice must learn to separate the wheat from the chaff, and perhaps the author relies too much for this on the classroom instructor. More essential is the question of which factors are to be neglected or crudely approximated. The answer must be a compromise between the puny powers of the human mind and the relevant physical magnitudes. To assay the latter requires experience or experiment. Professor Nikolsky appears to know his way around here. He knows what to discard and when to use a calculating machine, even depicting at one point an entry form and supplying a detailed example of its use.

The two chapters on stability are introduced by one on linear differential equations with constant coefficients. Even covered is such recurrent ground as the theories of characteristic roots (here termed "solution by differential calculus") and the Laplace transform ("Operational Calculus"). Why this instead of a reference to one of the numerous texts? On page 142, the author states that the problem of dynamic stability consists in solving a system of such equations. This is hardly correct. The criterion is that of Hurwitz-Routh, of which we could detect but a fleeting glimpse in one line. However, the chapter elucidates some lesser known and valuable numerical methods pertinent to the present field.

The style is like that of too many engineering texts—dull and unstimulating. This is partially the fault of custom and of editors who ban in works of this sort the indefinite pronoun—the English *we*, the French *on*, the German *man*. Although such are freely accepted in more theoretical tracts, a bizarre creed of artificiality cripples syntax in "applied" works. The author is bound to the passive voice. He may say "The equation is solved," but never "We solve the equation." Pernicious dullness results too often and we, *for one*, are tired of it.

RUFUS ISAACS
The Rand Corporation

Economic Aspects of Atomic Power

BY SAM H. SCHURR AND JACOB MARSCHAK. Pp. 289+xxvi. Princeton University Press, Princeton, New Jersey, 1950. Price $6.00.

The number of professional people engaged in the atomic energy field has rapidly expanded in the past five years. There are considerable numbers who hardly consider themselves in the field but are actively contributing to one or another phase of some Atomic Energy Project, and there is every indication that the number of engineers, chemists, production men in all types of industrial fields, management personnel, and professional consultants having interests and responsibilities in such projects will steadily increase.

Such participation brings with it the demand for explanation, for evaluation, and for prediction. To remain within the bounds of security regulations and still cover the need with authority, with plausible and well-founded speculation, and with caution rather than either popular enthusiasm or easily accomplished pessimism, became a project for Sam H. Schurr and Jacob Marschak.

It is unfortunate that it cannot be assumed that the technical capabilities and characteristics of the fission process are generally understood by the average professional reader. This basic education of the reader is handled in Part One of "Economic Aspects of Atomic Power" without much concession to the inquiring mind except to furnish an excellent bibliography of currently available unclassified information. No time or words are lost in popularized or story-like presentation of the fission process or basic fission physics. This is a relief.

The assumptions made have plausible technical defense although insufficient evidence is supplied to allow the reader to draw conclusions from basic facts. The entire fuel handling and chemical processing is hardly given a proportionate treatment, but the reader is assured that such cost factors have been duly considered.

The authors have chosen to apply the presently known published information about a new energy source quite rigorously to established industrial needs. This is an extremely practical and well-accomplished approach. Their presentation of the cost elements and an analysis of the effects of changes in economics of an industry as influenced by the character and cost of the prime energy is informative and revealing and should be of considerable interest for its own sake.

A basically new process, a new material, a radically new method of energy conversion, or as in this case, a new source of energy, suffers by application or evaluation on the basis of established needs. New discoveries have in the past developed their own market and application to a much higher degree of efficiency than merely substitution in a current process, industry, or use. It is difficult with the restricted knowledge permitted for general public appraisal to speculate intelligently on the possible revolutionary applications which may be found. The authors' approach has successfully offered opportunity for such reader speculation, however; and to their credit they have widened the scope of uses to include an intelligent discussion of the direct use of heat energy without a conversion to electric power.

The conclusions drawn are noncommittal. The authors create an impression of unwillingness to make a positive judgment but rather find that in all phases of application considered the new energy source will probably be competitive but not markedly so. This is safe, but not too satisfying.

For its time the survey is well done, authoritative, and of value and interest not only to professional people in the open field of industrial engineering but also to all classes within the atomic energy field itself. I have frequently recommended its merit to visitors from utilities and interested consulting firms.

K. A. KESSELRING
General Electric Company

Radioactivity Applied to Chemistry

ARTHUR C. WAHL, *Editor*, AND NORMAN A. BONNER, *Assistant Editor*. Pp. 604. John Wiley & Sons, Inc., New York, 1951. Price $7.50.

The book presents an exhaustive and well-documented literature search on the *application* of radioactivity to chemistry. Part I (320 pp.) consists of 10 chapters dealing with the following topics: isotopic exchange reactions, chemical kinetics, structural chemistry, self-diffusion, analytical applications, coprecipitation phenomena, discovery and properties of the newer elements, "hot-atom chemistry," emanation (from solids) methods, and specific surface determination. The literature is complete to January, 1950.

Although the entire book has been written by 12 contributors, there is a remarkable uniformity in the presenattion of a given subject in each chapter. The treatment is clear, systematic, adequate, and in some instances critical in nature. A wealth of information of interest to all chemists and especially to physical, inorganic, and analytical chemists is given. The preface states: "This book has been written for the reader with a good knowledge of chemistry but not necessarily more than an elementary knowledge of radioactivity. It is intended primarily as a reference book . . . ." It seems admirably suitable as a reference book for seminars in the various fields of chemistry for an up-to-date discussion of problems discussed in the text. Since the major emphasis is on the use of radioactive methods, other sources should also be consulted for a complete discussion of a particular subject. As an illustration, in a brief discussion on the use of radioactivity in studies of reactivity of the solid state no mention is made of Hedvall's extensive (nonradioactive) work in this field. Similarly, chapter 10, dealing with the determination of specific surface, is limited to radioactive methods. Chapter 6, dealing with coprecipi-

tation phenomena concentrates on the work of Otto Hahn and his school. Unfortunately, the authors of this chapter have copied Hahn's statement: "When incorporation into the lattice occurs, the ratio of the fraction of tracer to the fraction of carrier removed from solution is relatively independent of precipitating conditions such as the presence or absence of an excess of the precipitant or of highly charged ions." This would be true only if distribution equilibrium were obtained or approached during precipitation, which is usually not the case, and which is also apparent from the contents of the chapter.

The above comments do not detract from the usefulness of the book for the purposes for which it is written. It makes interesting reading to all chemists and should be especially stimulating to graduate students for further study and research.

Part II consists of eighteen extensive tables furnishing a summary of and a key to the literature of the various applications of radioactivity to chemistry. The preparation of these tables must have involved a great deal of time and work and will be a source of frequent consultation by workers on the subjects discussed in the book.

Altogether, the book is a welcome and valuable addition to the chemical literature.

I. M. KOLTHOFF
University of Minnesota

Other Books Received

Technical Publications 1949. ARMAN A. BECKER, *Editor*. Pp. 563+xi. Standard Oil Company (New Jersey), New York, 1952.

Manufacturing Processes. MYRON L. BEGEMAN. Third Edition. Pp. 608+viii. John Wiley & Sons, Inc., New York, 1952. Price $6.00.

Magnetic Materials. F. BRAILSFORD. Second Edition. Pp. 156. John Wiley & Sons, Inc., New York, 1951. Price $1.50.

Servomechanisms and Regulating System Design. HAROLD CHESTNUT AND ROBERT W. MAYER. Vol. I. Pp. 505+xiii. John Wiley & Sons, Inc., New York, 1951. Price $7.75.

French-English Science Dictionary. LOUIS DEVRIES. New Second Edition. Pp. 596+xii. McGraw-Hill Book Company, Inc., New York, 1952. Price $6.50.

Fundamentals of Automatic Control. G. H. FARRINGTON. Pp. 285+168. John Wiley & Sons, Inc., New York, 1951. Price $5.00.

Electrotechnics. B. L. GOODLET. Pp. 247+viii. Longmans, Green and Company, New York, 1951. Price $4.00.

Linear Transformations in *n*-Dimensional Vector Space. H. L. HAMBURGER AND M. E. GRIMSHAW. Pp. 195+vii. Cambridge University Press, New York, 1952. Price $4.50.

Principles of Radio. KEITH HENNEY AND GLEN A. RICHARDSON. Sixth Edition. Pp. 655+vii. John Wiley & Sons, Inc., New York, 1952. Price $5.50.

An Introduction to Electron Optics. L. JACOB. Pp. 48. John Wiley & Sons, Inc., New York. 1951. Price $2.00.

TV and Electronics as a Career. IRA KAMEN AND RICHARD H. DORF. Pp. 326. Figs. 136. John F Rider, Publisher, New York, 1951. Price $4.95.

D-C Power Systems for Aircraft. R. H. KAUFMANN AND H. J. FINISON. Pp. 206+x. John Wiley & Sons, Inc., New York, 1952. Price $5.00.

Dimensional Analysis and Theory of Models. HENRY L. LANGHAAR. Pp. 166+xi, Figs. 15. John Wiley & Sons, Inc., New York, 1951. Price $4.00.

Absorptions-Spektralanalyse. FRANZ X. MAYER AND ALFRED LUSZCAK. Pp. 238+xiv. Figs. 74. Walter de Gruyer and Company, Berlin, 1951.

Hochfrequenztechnik und Weltraumfahrt. R. MERTEN. Pp. 116, Figs. 65. S. Hirzel Verlag, Berlin, 1951.

Transients in Power Systems. HAROLD A. PETERSON. Pp. 361+xi. John Wiley & Sons, Inc., New York, 1951. Price $6.50.

Radiations from Radioactive Substances. ERNEST RUTHERFORD, JAMES CHADWICK, AND C. D. ELLIS. Pp. 588+xi, Figs. 140. Cambridge University Press, New York, 1952. Price $11.00.

Essay in Physics. HERBERT L. SAMUEL. Pp. 178+x. Harcourt, Brace and Company, Inc., New York, 1952. Price $3.00.

Fourier Transforms. IAN N. SNEDDON. Pp. 542+xii, Figs. 108. McGraw-Hill Book Company, Inc., New York, 1951. Price $10.00.

Tables of the Error Function and of Its First Twenty Derivatives. Staff of the Computation Laboratory of Harvard University. Pp. 276+xx. Harvard University Press, Cambridge, Massachusetts, 1951. Price $8.00.

College Physics. ROBERT L. WEBER, MARSH W. WHITE, AND KENNETH V. MANNING. Second Edition. Pp. 820+viii. McGraw-Hill Book Company, Inc., New York, 1952. Price $6.50.